

# SEQUENCE LISTING

5           <110> University of Iowa Research Foundation et al.  
           <120> Novel polypeptides and methods of their use  
           <130> 875.006WO1  
 10          <150> US 60/105,575  
           <151> 1998-10-26  
           <160> 65  
 15          <170> FastSEQ for Windows Version 3.0  
           <210> 1  
           <211> 2265  
           <212> DNA  
 20          <213> Artificial Sequence  
           <220>  
           <223> A modified DNA sequence  
 25          <400> 1  
           gagggtagag aaatcgaatg ccccttgcac caaggtcggg ttgacgtttg cacaggcaaa       60  
           gccctgtgcg caccctgac acagaacatc aaaacatcag cagtcaagat tgagaacctg       120  
           cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt       180  
           aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggg       240  
 30          aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca       300  
           tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc       360  
           tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga       420  
           gaacgacggg tcgattcgtg cttttctgaa cgtttgccg catcgtggca agacgtggg       480  
           gagcgtggaa gccggcaatg ccaaaggttt tgtttcgag tatcacggc ggggcttcgg       540  
 35          ctccaacggg gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa       600  
           taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta       660  
           cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg       720  
           cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt       780  
           tgtgatcaag gccaaactgga aggcaccgcg ggaaaacttt gtgggagatg cataccacgt       840  
 40          gggttggaag cagcgtctt cgttcgctc gggggagtct atcttctcgt cgtcgtcgtg       900  
           caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg       960

```

catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020
gatggcatcc ggaggcgcaa agcaggaaaag gctgaacaaa gaaattggcg atgttcgcgc 1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
ctcgggtgtt ttcaaagtat ggaaccgat cgacgcaaacc accaccgagg tctggacctt 1200
5 cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca 1260
gcgaacggtc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttgggtt 1380
cgggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaact cggcgatcgg 1440
cgagaccagt tategtggtt tctaccgggc ttaccaggca cagtcagca gctccaactg 1500
10 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggg ttccgcccac 1620
gacgcgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgaccagga agcgcatctg ttggacatcc aggettaccg tgcttgggta 1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cgcgcgaact gcgcgcagct 1800
15 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagtcc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaactg cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
20 cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
tgattcagtg accattttta caaatgggta ctgcaaccgc ggtcaccatt aatcaaaggg 2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

```

<210> 2

25 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

30 <223> A polypeptide encoded by SEQ ID NO:1

<400> 2

```

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
  1             5             10             15
35 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
      20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
      35             40             45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
40      50             55             60

```

Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65                      70                      75                      80  
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
                     85                      90                      95  
 5 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
                     100                      105                      110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
                     115                      120                      125  
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 10      130                      135                      140  
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145                      150                      155                      160  
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
                     165                      170                      175  
 15 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
                     180                      185                      190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
                     195                      200                      205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
 20      210                      215                      220  
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
 225                      230                      235                      240  
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
                     245                      250                      255  
 25 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
                     260                      265                      270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
                     275                      280                      285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 30      290                      295                      300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305                      310                      315                      320  
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
                     325                      330                      335  
 35 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val  
                     340                      345                      350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
                     355                      360                      365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 40      370                      375                      380

10

<211> 9841

15

<213> Artificial Sequence

<223> A modified DNA sequence

<400> 3

	gaattcatca	ggaagacatt	caaatgaacg	taaacaataa	gggcagcgtc	tgtatttgcg	60
	gcagcgaaat	gctccctaaa	ttctcattt	accccatctg	aggattgctt	tatgacagta	120
	aagtggattg	aagcagtcgc	tctttctgac	atccttgaag	gtgacgtcct	cggcgtgact	180
	gtcggaggga	aggagctggc	gctgtatgaa	gttgaaggcg	aaatctacgc	taccgacaac	240
25	ctgtgcacgc	atgggtccgc	ccgcatgagt	gatggttatt	tcgagggtag	agaaatcgaa	300
	tgccccttgc	atcaaggctg	gtttgacgtt	tgcacaggca	aagccctgtg	cgcaccctgt	360
	acacagaaca	tcaaaacata	tccagtcaag	attgagaacc	tgcgcgtaat	gattgatttg	420
	agctaagaat	tttaacagga	ggcaccctcg	gccctagagc	gtaatcacc	ccattccatc	480
	tttttttaggt	gaaaacatga	attacaataa	taaaaatctt	gtaagtgaat	ctggtctgag	540
30	ccaaaagcac	ctgattcatg	gcgatgaaga	acttttccaa	catgaactga	aaaccatttt	600
	tgcgcggaac	tggctttttc	tactcatga	tagcctgatt	cctgccccc	gcgactatgt	660
	taccgcaaaa	atggggattg	acgaggtcac	cgtctcccgg	cagaacgacg	gttcgattcg	720
	tgtttttctg	aacgttttgc	ggcatcgtgg	caagacgctg	gtgagcgtgg	aagccggcaa	780
	tgccaaaggt	tttgtttgca	gctatcacgg	ctggggcttc	ggctccaacg	gtgaactgca	840
35	gagcgttcca	tttgaaaaag	atctgtacgg	cgaagtcgctc	aataaaaaat	gtctgggggt	900
	gaaagaagtc	gctcgcgtgg	agagcttcca	tggcttcac	tacggttgct	tcgaccagga	960
	ggccctcct	cttatggact	atctgggtga	cgtcgttg	tacctggaac	ctatgttcaa	1020
	gcattccggc	ggtttagaac	tggtcggtc	tccaggcaag	gttgtgatca	aggccaactg	1080
	gaaggcacc	gcggaaaact	ttgtgggaga	tgcataccac	gtgggttgga	cgcacgcgtc	1140
40	ttcgcttcgc	tcgggggagt	ctatcttctc	gtcgtctcgt	ggcaatgcgg	cgctaccacc	1200

tgaaggcgca	ggcttgcaaa	tgacctccaa	atacggcagc	ggcatgggtg	tgttggtgga	1260
cggatattca	ggtgtgcata	gcgcagactt	ggttcggaa	ttgatggcat	tcggaggcgc	1320
aaagcaggaa	aggctgaaca	aagaaattgg	cgatgttcgc	gctcggattt	atcgcagcca	1380
cctcaactgc	accgttttcc	cgaacaacag	catgctgacc	tgctcgggtg	ttttcaaagt	1440
5atggaacccg	atcgacgcaa	acaccaccga	ggtctggacc	tacgccattg	tcgaaaaaga	1500
catgcctgag	gatctcaagc	gccgcttggc	cgactctgtt	cagcgaacgg	tcgggcctgc	1560
tggtctctgg	gaaagcgacg	acaatgacaa	tatggaacaa	gcttcgcaaa	acggcaagaa	1620
atatcaatca	agagatagtg	atctgctttc	aaaccttggg	ttcggtgagg	acgtatacgg	1680
cgacgcggtc	tatccaggcg	tcgtcggcaa	atcggcgatc	ggcgagacca	gttatcgtgg	1740
10tttctaccgg	gcttaccagg	cacacgtcag	cagctccaac	tgggctgagt	tcgagcatgc	1800
ctctagtact	tggcatactg	aacttacgaa	gactactgat	cgctaacaga	cgagtcgacc	1860
atgatgatca	atattcaaga	agacaagctg	gtttccgccc	acgacgccga	agagattcct	1920
cgtttcttca	attgccacga	ctctgctttg	caacaagaag	ccactacgct	gctgaccacg	1980
gaagcgcat	tgttgacat	tcaggcttac	cgtgcttggg	tagagcactg	cgtgggggtca	2040
15gaggtgcaat	atcaggtcat	ttcacgcgaa	ctgcgcgcag	cttcagagcg	tcgttataag	2100
ctcaatgaag	ccatgaacgt	ttacaacgaa	aattttcagc	aactgaaagt	tcgagttgag	2160
catcaactgg	atccgcaaaa	ctggggcaac	agcccgaagc	tgcgctttac	tcgctttatc	2220
accaacgtcc	aggccgcaat	ggacgtaaat	gacaaaagagc	tacttcacat	ccgctccaac	2280
gtcattctgc	accgggcacg	acgtggcaat	caggctcgatg	tcttctacgc	cgcccgggaa	2340
20gataaatgga	aacgtggcga	agggtggagta	cgaaaattgg	tccagcgatt	cgtcgattac	2400
ccagagcgca	tacttcagac	gcacaatctg	atggctcttc	tgtgattcag	tgaccatttt	2460
tacaaatggg	cactgcaacc	gcggtcacca	ttaatcaaag	ggaatgtacg	tgtatgggca	2520
atcaacaagt	cgtttcgata	accggtgcag	gctcaggaat	cggtctcgaa	ctgggttcggg	2580
cctttaagtc	ggccggttat	tacgtatccg	ctctcgtacg	aaacgaggag	caagaggcgc	2640
25ttctttgcaa	agagttcaag	gacgcactcg	agattgtagt	gggcgatgtc	cgggaccacg	2700
caacaaatga	gaagctgata	aagcaaacaa	tcgatagatt	cggtcactct	gattgtttta	2760
ttgcaaatgc	cggatatctg	gattacatgc	tgagcatcga	agagccttgg	gagaaaaat	2820
cgagcagttt	tgacgaaata	ttcgacatta	atgtcaagag	ctatttcagt	ggcatcagtg	2880
ccgccctgcc	ggaaactgaaa	aagactaacg	gatcagtggt	gatgaccgct	tcggtgtcgt	2940
30cccatgcggt	cgggtggtgg	ggttcttgct	acatcgccag	caagcatgcg	gtgctcggta	3000
tggttaaggc	tttggcctac	gaattggccc	ccgaagtctg	cgtgaacgct	gtttcgccgg	3060
ggggcaccgt	gacgtctctg	tgcggtcccg	cgagcgccgg	tttcgacaaa	atgcacatga	3120
aagacatgcc	cggcatcgac	gatatgatca	aaggtctcac	gcctcttggg	tttgacagcca	3180
agcccgaaga	cgtggtggca	ccctatttgt	tgctggcttc	gcgaaagcaa	ggaaaattca	3240
35tcaccggcac	cgtgattagc	attgatggcg	gtatggcgct	cggtcgcaag	tgagcttgta	3300
gccgatcaga	agttatagac	acatttcagg	tgacgcccc	tgaagacaaa	actgtttatc	3360
aataacgcct	ggatcgattc	tagtgaccag	cagaccttcg	agcgcataca	ccccgtcagc	3420
agcgatgtgg	tgactgagag	cgcaaacgcc	acagtgcagg	acgcgataaa	ggcggcgcaa	3480
gcggccgagg	aggcgttcaa	gacctggaag	gccgttgac	cttcagagcg	tcgccgcctt	3540
40ctcctaagg	tcgccgatgt	catggaaagt	aaaacacca	agttcatcga	agtgatggcc	3600

	atggaggttg	gagcttccgc	cctttgggcc	ggatttcaacg	tccatgcgtc	tgccaatgtg	3660
	ttccgagagg	ctgcctcgc	ggctacccaa	attcagggtg	aaaccatccc	aacggacaaa	3720
	gccgaaacgc	tctcaatgac	actacgtcag	ccggtcggcc	cgatcctaag	catcgttcca	3780
	tggaacggca	ccgcagtgct	tgcggcacga	gccatcgctt	atccgctggg	ctgtggcaac	3840
5	actgtggtgt	tcaaaggctc	tgaatttagt	cccgcgacgc	atgccctgat	caccagtg	3900
	gtgcaggaag	ccgggctgcc	cgctggcgtg	ctcaattacc	tcaactcttc	gcctgaccgt	3960
	tgcgccgaga	tgcgtgacgc	actgatctct	gccaaaggaga	tccgcgcgat	caacttcacg	4020
	ggttccacc	gcgtgggcag	cattatcgcg	cagaaagccg	cgcaacacct	caagcgtctg	4080
	ctgctggagc	tccggggcaa	gtccccgctt	attgttctgg	atgatgcaga	catcgatgcg	4140
10	gcggtcaagg	cagcgggtgt	cggtagcttc	ctgttccaag	gtcagatctg	catgtccact	4200
	gagcgcctga	tcgttgatga	gaagatagcc	gacgaatttg	tgcgaaaatt	tgtcgaaaaa	4260
	actaagcgct	tgagcgcagg	cgacccgtgc	gtaactggcg	actgcatcat	cggcccgatg	4320
	gtctcgccaa	attcgggtga	gcggatcaat	ggtttgttca	aagacgcgat	cgacaaaggg	4380
	gcaaaagttg	tttgcggcgg	cttgccccaa	gggtgcgtca	tgccggccac	gatcctggat	4440
15	cacgtcaaat	ctgacatgcg	gatttacgat	gaggagacct	ttgggtcccat	cacggtggt	4500
	atccgttgta	aaggcgaagc	agaggccgtc	cgcattgcc	acgacagcgt	ctatggcctg	4560
	tgcgcggcg	tatttgccg	cgacatcaac	cgcgtctctac	gcgtgggtat	gtccatcgaa	4620
	tatggttctg	tacacatcaa	cggttcgacc	gtccagaacg	aggcgcaggc	tccttacgga	4680
	ggcaccaaga	acaccggcta	cgggcgcttc	gacggccgtg	ctgtaatcga	cgagttcaca	4740
20	gagatcaagt	ggctgaccat	cgaacctttc	gagcagcaat	atcccttctg	ataagcacta	4800
	actccaggga	atcaaactat	gagtaagcaa	gctgcagtta	tgcagctcgg	atacatgggt	4860
	atctcgggtca	aggacctga	tgcgtggaaa	tcatttgcca	cggatatgct	agggtctgcaa	4920
	gttcttgatg	agggtgagaa	ggaccgtttc	tatctgcgga	tggattactg	gcatcatcgg	4980
	atcgtagtcc	atcacaacgg	acaggacgac	ttggagtacc	taggctggcg	tgtagccggc	5040
25	aagccggagt	tccaagctct	gggtcaaaaag	cttattgatg	ccggttacaa	gatccgcata	5100
	tgcgacaaa	ttgaggtctc	ggagcgtatg	gtgttgggtc	tgatgaagac	agaagatccg	5160
	ggcggcaacc	cgaccgagat	attctggggc	ccccggatcg	acatgagcaa	cccgttccat	5220
	cccggtcgcc	ccctgcacgg	aaagtttgtg	accggtgacc	aaggcttggg	ccattgcata	5280
	gttcgccaaa	ccgacgtcgc	agaagctcat	aagttttata	gcctgctggg	cttcctgggg	5340
30	gacgtcgaat	accggtatcc	gttgcccaac	ggcatgactg	ccgaactgtc	gttcatgcat	5400
	tgcaacgccc	gtgatcactc	cattgtcttt	gggtccatgc	ccgtgcgcaa	acgactcaat	5460
	cacttgatgc	ttgagtacac	ccatatggaa	gacttgggat	acacgcacca	acagtttgta	5520
	aagaacgaaa	ttgacattgc	cttgacgctt	ggcattcacg	ccaacgacaa	ggcgttgacg	5580
	ttctatgggt	caacgccttc	gggctggctc	attgagcccc	gctggcgagg	tgccacggcc	5640
35	atagatgaag	cggagtatta	cgtcggcgac	atcttcggcc	atggcgtgga	ggccactgga	5700
	tatggccttg	atgtaaaact	gagctaaaga	tgcgcgctcg	ttgggcgagg	ctctagtcca	5760
	gcatcttc	acgcaaccaa	ccttgacagg	cgatgagatc	aaaggacgtt	aaagcgaagg	5820
	ggaagtgggt	cgggccatgc	gcataccgat	ccatgacatt	tgtttcatag	tatataggta	5880
	gatagggtgaa	tcaagcgctt	agtcaactag	tggacacatc	tgttccatga	ggctatctac	5940
40	tatctattca	aaacaagaat	aataaatagg	atgaaaataa	taatgataaa	aagaacgatt	6000

tgtcttgtgt atcctctatt ctgtttggca agccccacat gggccgaaga gtgcgcttgg 6060  
 acgtaccgta ttggtatgac taatgtagct ttcgatgcta gcgcaaaagt atacttaaat 6120  
 ggtcagcggg tgccaggagg aagcgtgat gcgagcgata acaacgcgct tacattcgac 6180  
 ttcggtctacg ccatcaacga ccagtggaa gtacgtgcga ttgtcggtat tccgcctaca 6240  
 5actaaagtga cgggcgcagg cacttctct ggtatccagc tggggaaaat aacttacgct 6300  
 ccaacagtat taacgttgaa ctataacctc cccgctttgg gtcccgctcg cctcacata 6360  
 ggtgcgggag tcaattacac gcggattttt gaaagtcggg acgctaattct aaaatcgttc 6420  
 gatgccgacc acgcttggtc ccccgcgcta catgttggtg ccgatatga cgttaaccgt 6480  
 ggttggttcg ttagcattga tatccggaag ttataacctga aaaccgacgc atcagggtag 6540  
 10ttggggccac aggaggctaa agcacgggta actcttgacc cattactaac ttcgatcgcg 6600  
 atcggacgcc aattctgatg attctgttta agttcttta tctatctaac cgcaaaggg 6660  
 gtttccatgt cgaataaaat tatgaaaacg tcgctctta ccgccgaaga tatcaacggc 6720  
 gcctggacta taatgccac accctcgacg cctgatgctt ctgattggcg cagcactgcc 6780  
 accgtggact tagaagagac tgcccgcata gttgaagagc tgattgcagc tgggtgtcaac 6840  
 15ggtattctaa gtatgggtac ttttgggtgag tgcgccacgt tgacctggga tgaaaaacgt 6900  
 gattatgtct cgacgattgt cgagaccatt cgtggtcgcg tgccttattt ctgtggcacg 6960  
 acagccttaa ataccgaga agtcatccgc cagaccgag agcttatcga tattggcgcc 7020  
 aacggcacta tgctcggggt gccgatgtgg gtgaagatgg acctgcctac agcggttcag 7080  
 ttctatcgtg atgttgaga tgcggtacca gaggtgccca ttgcgattta cgccaaaccc 7140  
 20gaagcattca agttcgactt cctcgccca ttctgggcag agatgtccaa aattccgcag 7200  
 gtagtgactg cgaagtatct aggcacgga atgcttgact tggacctgag actggcacc 7260  
 aacatccgct tcttcccca cgaagatgac tattacggcg ccgcacgcat caatcccgag 7320  
 cgcataaccg cgttctggtc aagcggggcc atgtgcggcc cggtaccgc catcatgttg 7380  
 cgtgacgaag tgggtcgggc caagagcacc ggtgactggg ccaaggccaa agccatctcc 7440  
 25gatgatatgc gtgcagccga ctcgacattg tttccgctg gcgacttttc ggagtctctg 7500  
 aagtataata tcgggcttga aaaggcagcg atggatgcgg ctggttggct caaggctggg 7560  
 cctgcgctc cgccctacaa ccttgttcca gaagactacc tcgctggtgc acagaaatca 7620  
 ggcaaggctt gggccgctg gacgctaaa tacagtaatg aattgaagta gttcacctcc 7680  
 gcagacctga gtgacagggt ggcgcagacg ctgaggggtgc aggaattaag tgagctaaag 7740  
 30cacatttctt gcgccaggca ttgccagatc agcaaagttt gctgatctgg cagtttcaaa 7800  
 aatttgggcg aaagctgata tcaggaatac gggataaagg cagtgcacca taacgacggg 7860  
 gcgtgccatt cgtgatgaac gatcttgcta ttgtgccgac ttctgttctt ggagtgttg 7920  
 attgtgattg tcgattttta tttcgatttt ttgagtccgt tctcttactt ggccaacct 7980  
 cgtttgtcaa agcttgcgca agactatggc ttttccattc gttattacgc aatcgatttg 8040  
 35gcgcgagtta aaatagccat cggaaacgtt ggtccatcta atcgcgacct gatagtcaag 8100  
 ctggactatt tgaaagtaga tttgcaacgg tgggccgagc tttacgaaat accgttggt 8160  
 ttccagcta actacaacg ccgacggatg aatactgggc tttattactc gggagccatg 8220  
 gcacagactg gtgcctatgt gaatgtagta tttaatgcgg tttggggaga tggcatagct 8280  
 ccagatttgg aaagcttgcc tgctctggta tctgaaaaac taggctggga tcgtagcgcc 8340  
 40ttcgaggact ttatcagcag cgatgccgca acagagaggt atgacgagca gacacatgcc 8400

gcgatcgaac gcaaagtgtt cgggtgtgcca acgatgtttt tgggcatga aatgtggtgg 8460  
 ggaaacgacc gtctatttat gctcgagaac gcagtgggag gtgcgcctgt aaatggagaa 8520  
 tagtcgctac ggagcgcttg tgcgggctaa atgccgatat aagtgggtga cctgatcggt 8580  
 atttgctcga tacagcgctt tcaaaatcag cggctactga agtcagataa aaatgcggga 8640  
 5ctacttcagg catcctgtgc gacacaaagt ttacctgta attgtccacc tattccgagt 8700  
 ttggaatggg agctgactcg ctatgcgacc agcgatagcc taacaagaca tgcactactg 8760  
 gtaacgggtgg ggtgtgaagc tcctgcaaca atgtagcccc ttgatgtgtg tatttgcctgc 8820  
 gaggtgaagc acagatgctc ggagcgctac cggcttctgg cgctaggctg gcaagtatga 8880  
 gcaacgtaag tgggggttgg ggcgcaatgg gaacaaaaa ccaacgcaag ccttaccagc 8940  
 10gtcgttcggt gccttcctcc catgcctccg cctcgataaa gcagctgcgc atatcggtct 9000  
 cctggctgat ctcggttagt aggtcatgca aggtcttctc cagcgctcgc tcgctccgat 9060  
 acggaatggg cagctcgtaa tggcgggtct ccgaccgctt catgcgctag ggctccaggc 9120  
 agtagcgctc gatgttctcc gtggcccgct tccgaccgag catgaacttg ctgttggtca 9180  
 ccaccgccag gcgcagggtg acggtggcca cccgctcgac ggttgactct gccggtgacg 9240  
 15cgatattgct cttttgacct cgcgccaggg cgctctctct gtacgtcccg atctcgacgc 9300  
 cacgggtggc taggtagctg tacagggtgc tcttgagat gtgcaactc tcgccgatgg 9360  
 cgctgacgct caggcgcccc tcgcggtaca gggctctcgc cgccatggcg gtggcctcgg 9420  
 ccttggtcgg caggcccttg ggacggcgac cgatccggcc tcgagtcctg gccgccgaca 9480  
 ggcccgctcg agtcgctcg cggatcagct cgcgctcgaa ctgaacaggt tgaacaccag 9540  
 20gcgatcttgg gcgtgggtgc tgtcaatggg gtgcttcagg ctctgcaagc cgactttgct 9600  
 tgcagccagc tagccgacca actcaaccag gtgcttgagc gagcgaccga ggcgatccag 9660  
 ctccagatc accacggcat cgcgcgctcg aacatgggct agcaacttgt ccaactccgg 9720  
 ccgcgcgctt tttgcgcgcg tggcgatgct ttgatagatg cgttcgcacc cggcctgttt 9780  
 cagggcacgc acctggaggt cggcgcttcta atcccagtg ctcaccgcgc tataaccgat 9840  
 25c 9841

<210> 4

30 <211> 2515

<212> DNA

<213> Artificial Sequence

<220>

35 <223> A modified DNA sequence

<400> 4

gaattcatca ggaagacatt caaatgaacg taaacaataa gggcagcgctc tgtatttgct 60  
 gcagcgaaat gctccctaaa ttctcatatt accccatctg aggattgctt tatgacagta 120  
 40 aagtggattg aagcagtcgc tctttctgac atccttgaag gtgacgtcct cggcgtgact 180



	gtcagaggca	aggagctggc	gctgtatgaa	gttgaaggcg	aaatctacgc	taccgacaac	240
	ctgtgcacgc	atggttccgc	ccgcatgagt	gatggttata	tcgagggtag	agaaatcgaa	300
	tgcccccttg	atcaaggctg	gtttgacgtt	tgcacaggca	aagccctgtg	cgcacccgtg	360
	acacagaaca	tcaaaacata	tccagtcaag	attgagaacc	tgcgcgtaat	gattgatttg	420
5	agctaagaat	tttaacagga	ggcaccgccg	gccctagagc	gtaatcaccc	ccattccatc	480
	tttttttaggt	gaaaacatga	attacaataa	taaaatcttg	gtaagtgaat	ctggtctgag	540
	ccaaaagcac	ctgattcatg	gcgatgaaga	acttttccaa	catgaactga	aaaccatttt	600
	tgcgcggaac	tggctttttc	tactcatga	tagcctgatt	cctgcccccg	gcgactatgt	660
	taccgcaaaa	atggggattg	acgaggtcat	cgtctcccg	cagaacgacg	gttcgattcg	720
10	tgcttttctg	aacgtttgcc	ggcatcgtgg	caagacgctg	gtgagcgtgg	aagccggcaa	780
	tgccaaaggt	tttgtttgca	gctatcacgg	ctggggcttc	ggctccaacg	gtgaactgca	840
	gagcgttcca	tttgaaaaag	atctgtacgg	cgagtcgctc	aataaaaaat	gtctgggggt	900
	gaaagaagtc	gctcgcgtgg	agagcttcca	tggcttcac	tacggttgct	tcgaccagga	960
	ggccctcct	cttatggact	atctgggtga	cgctgcttgg	tacctggaac	ctatgttcaa	1020
15	gcattccggc	ggtttagaac	tggtcggctc	tccaggcaag	gttgtgatca	aggccaactg	1080
	gaaggcacc	gcggaaaact	ttgtgggaga	tgcataccac	gtgggttgga	cgcacgcgtc	1140
	ttcgtctcgc	tcgggggagt	ctatcttctc	gtcgtcgcgt	ggcaatgcgg	cgtaccacc	1200
	tgaaggcgca	ggcttgcaaa	tgacctccaa	atacggcagc	ggcatgggtg	tgttggtgga	1260
	cggatattca	ggtgtgcata	gcgcagactt	ggttccggaa	ttgatggcat	tcggaggcgc	1320
20	aaagcaggaa	aggctgaaca	aagaaattgg	cgatgttcgc	gctcggattt	atcgcagcca	1380
	cctcaactgc	accgttttcc	cgaacaacag	catgctgacc	tgctcgggtg	ttttcaaagt	1440
	atggaacccg	atcgacgcaa	acaccaccga	ggtctggacc	tacgccattg	tcgaaaaaga	1500
	catgcctgag	gatctcaagc	gccgcttggc	cgactctggt	cagcgaacgg	tcgggcctgc	1560
	tggcttctgg	gaaagcgacg	acaatgacaa	tatggaaaca	gcttcgcaaa	acggcaagaa	1620
25	atatcaatca	agagatagtg	atctgctttc	aaaccttggg	ttcggtgagg	acgtatacgg	1680
	cgaacgcggc	tatccaggcg	tcgtcggcaa	atcggcgatc	ggcgagacca	gttatcgtgg	1740
	tttctaccgg	gcttaccagg	cacacgtcag	cagctccaac	tgggctgagt	tcgagcatgc	1800
	ctctagtact	tggcatactg	aacttacgaa	gactactgat	cgctaacaga	cgaagtcgac	1860
	atgatgatca	atattcaaga	agacaagctg	gtttccgccc	acgacgccga	agagattctt	1920
30	cgtttcttca	attgccacga	ctctgctttg	caacaagaag	ccactacgct	gctgaccacg	1980
	gaagcgcatt	tgttgacat	tcaggcttac	cgtgcttggg	tagagcactg	cgtgggggtca	2040
	gaggtgcaat	atcaggtcat	ttcacgcgaa	ctgcgcgcag	cttcagagcg	tcgttataag	2100
	ctcaatgaag	ccatgaacgt	ttacaacgaa	aattttcagc	aactgaaagt	tcgagttgag	2160
	catcaactgg	atccgcaaaa	ctggggcaac	agcccgaaag	tgcgctttac	tcgctttatc	2220
35	accaacgtcc	aggccgcaat	ggacgtaaat	gacaaagagc	tacttcacat	ccgtccaac	2280
	gtcattctgc	accgggcacg	acgtggcaat	caggtcgatg	tcttctacgc	cgcgggggaa	2340
	gataaatgga	aacgtggcga	aggtggagta	cgaaaattgg	tccagcgatt	cgtcgattac	2400
	ccagagcgca	tacttcagac	gcacaatctg	atggtctttc	tgtgattcag	tgaccatttt	2460
	tacaaatggt	cactgcaacc	gcggtcacca	ttaatcaaag	ggaatgtacg	tgtat	2515

&lt;210&gt; 5

&lt;211&gt; 9706.

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

5

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

&lt;400&gt; 5

10gctgatttcg accaattatt ccacgccgtc tgcgccagca aggccgtaat cagttcgtaa	60
aggggagccc gcagtaacag cgaaccgacg ccggtagctg cgcgcgtgaa tcctgagatc	120
acaaattctt ggcgattatc gcatttcaat ttcacccaag cgcggttcct tgcgctgctg	180
tagcggcggc catcacgctt ctccacgatt ccctccagct gcacccggca agcgttgggc	240
agcagctgca ctgggtaagc ttccagggtg tcgctgaagc gtacgcgctc caatgggcag	300
15tgctcgagga ggcctcaag tacctggcgg ctgagttcaa cgccctctcc ccgcagatct	360
gcaccgtcga ggtagagcag gtcgaacgcc acgtacacaa atttgtcggg acaccgtgcg	420
gcgaacgctg actgcagggc ttgaaatgct ggtcgaccgt cgtcactact gaatacgacc	480
tcaccgtcga gccaggtcga atggacggct aggcgtccca gctatttggc caggagtggc	540
atatgatcga tctagtcaaa cccgtttttg gtgaaaagct gcacctggtc accatcgatc	600
20cgtgctaaca gacgatacca gtcgtacttg atctcgtatc gccatgctcc tgccgacgac	660
agtgtcagta gagtaaccag ttgtggcgcg atccaacacg ggtttggtgc cttgggcggg	720
ttcactatcg ccacctcgct ataggaacat tccttcaggg tggagtgcgt aattttctga	780
aaggggagcc aggttatgag tattcacatt ggtgataaac aacatcactt atgcgttatt	840
gacatataac gtcgtattca cgattattta ccatataagt cttataataa cgaagccata	900
25ttatggaact cctcatacaa ccgaaaaatc gcataattcc cttcagtgcc ggtgccaacc	960
ttctggaagt gcttcgcgag aacgggtgtag ctatttccta cagttgcttg tctgggcgtt	1020
gcggaacctg tcgctgccgg gttatagatg gcagtgtcat tgattctggg gcggaagt	1080
ggcaatcaaa cctcaccgac aagcagtatg tgctcgctg tcagtcagta ctactggca	1140
attgcgctat cgaagtcca gaagccgacg aaattgtcac tcaccggcg cgaatcatca	1200
30agggcacagt ggtcgcagtc gagtcgcca ctcacgatat ccgtcgctta cgcgtacgcc	1260
tctccaagcc cttcgagttc tcaccgggac agtacgcgac actgcagttc agccctgagc	1320
atgcgcgtcc gtattcaatg gcaggtttgc cagatgacca agaatggag ttccacatac	1380
gcaagggtgcc ggggtggcgc gtcacggagt atgttttcga acacgtccgc gaaggtaaa	1440
gcacaaagtt gagcgggcct cttgggtacgg cttatctacg tcagaagcac accggaccga	1500
35tgctgtgtgt aggtggcggg accggactcg caccgggtgt gtogattgtt cgcggcgcg	1560
tgaagtccgg tatgacgaac cccatcctcc tttatttcgg ggtgcgcagt cagcaagacc	1620
tctacgacgc agagcgattg caccaaactc ccgctgacca ccctcaactg accgtacaca	1680
cggtgattgc aacgggcccc attaatgagg gtcagcgagc cggcctaatt accgatgtga	1740
tcgaaaaaga catcctttcg ctggctgggt ggagggccta cctgtgcggc gcaccagcga	1800
40tggttgaagc gttgtgcacc gtcaccaagc atcttggaat atcaccgaa catatttatg	1860

	ccgatgcctt	ctatcccggt	gggatctgaa	tagttcccg	ccatgcacct	ctgtccatcg	1920
	agaattcatt	aggaagacat	tcaaatgaac	gtaaacaata	agggcagcgt	ctgtatttgc	1980
	ggcagcgaaa	tgtctccctaa	attcctcatt	taccccatct	gaggattgct	ttatgacagt	2040
	aaagtggatt	gaagcagtcg	ctctttctga	catecttgaa	ggtgacgtcc	tggcggtgac	2100
5	tgctcgagggc	aaggagctgg	cgctgtatga	agttgaaggc	gaaatctacg	ctaccgacaa	2160
	cctgtgcacg	catggttccg	cccgcacgag	tgatgggtat	ctcgagggta	gagaaatcga	2220
	atgccccttg	catcaaggtc	ggtttgacgt	ttgcacaggc	aaagccctgt	gcgcacccgt	2280
	gacacagaac	atcaaaacat	atccagtcga	gattgagaac	ctgcgcgtaa	tgattgattt	2340
	gagctaagaa	ttttaacagg	aggcaccocg	ggccctagag	cgtaatcacc	cccatcccat	2400
10	cttttttttagg	tgaaaacatg	aattacaata	ataaaatctt	ggtaagtga	tctggtctga	2460
	gccaaaagca	cctgattcat	ggcgatgaag	aacttttcca	acatgaactg	aaaaccattt	2520
	ttgcgcggaa	ctggcttttt	ctcactcatg	atagcctgat	tectgcccc	ggcgactatg	2580
	ttaccgcaaa	aatggggatt	gacgaggtca	tcgtctcccg	gcagaacgac	ggttcgattc	2640
	gtgcttttct	gaacgtttgc	cggcatcgtg	gcaagacgct	ggtgagcgtg	gaagccggca	2700
15	atgccc aaagg	ttttgtttgc	agctatcacg	gctggggctt	cggctccaac	ggtgaactgc	2760
	agagcgttcc	atgtgaaaaa	gatctgtacg	gcgagtcgct	caataaaaaa	tgtctggggg	2820
	tgaagaagt	cgctcgctg	gagagcttcc	atggcttcat	ctacggttgc	ttcgaccagg	2880
	aggccctcc	tcttatggac	tatctgggtg	acgctgcttg	gtacctggaa	cctatgttca	2940
	agcattccgg	cggtttagaa	ctggtcggtc	ctccaggcaa	ggttgatgac	aaggccaact	3000
20	ggaaggcacc	cgcgga aaac	tttgtgggag	atgcatacca	cgtgggttgg	acgcacgcgt	3060
	cttcgcttcg	ctcgggggag	tctatcttct	gctcgctcgc	tggcaatgcg	gcgctaccac	3120
	ctgaaggcgc	aggcttgcaa	atgacctcca	aatacggcag	cggcatgggt	gtgttgtggg	3180
	acggatattc	agggtgtgat	agcgcagact	tggttccgga	attgatggca	ttcgaggcgc	3240
	caaagcagga	aaggctgaac	aaagaaattg	gcgatgttcg	cgtcgggatt	tatcgcagcc	3300
25	acctcaactg	caccgttttc	ccgaacaaca	gcctgctgac	ctgctcgggt	gttttcaaag	3360
	tatggaaccc	gatcgacgca	aacaccaccg	aggctctggac	ctacgccatt	gtcgaaaaag	3420
	acatgcctga	ggatctcaag	cgccgcttgg	ccgactctgt	tcagcgaacg	gtcgggcctg	3480
	ctggcttctg	ggaaagcgac	gacaatgaca	atatggaaac	agcttcgcaa	aacggcaaga	3540
	aatatcaatc	aagagatagt	gatctgcttt	caaaccttgg	tttcggtgag	gacgtatacg	3600
30	gcgacgcggg	ctatccaggc	gtcgtcggca	aatcggcgat	cggcgagacc	agttatcgtg	3660
	gtttctaccg	ggcttaccag	gcacacgtca	gcagctccaa	ctgggctgag	ttcgagcatg	3720
	cctctagtac	ttggcatact	gaacttacga	agactactga	tcgctaacag	acgagtcgac	3780
	catgatgatc	aatattcaag	aagacaagct	ggtttccgcc	cacgacgccg	aagagattct	3840
	tcgtttcttc	aattgccacg	actctgcttt	gcaacaagaa	gccactacgc	tgctgaccca	3900
35	ggaagcgcac	ttgttgagca	ttcaggctta	ccgtgcttgg	ttagagcact	gcgtggggtc	3960
	agaggtgcaa	tatcaggtea	tttcacgcga	actgcgcgca	gcttcagagc	gtcgttataa	4020
	gctcaatgaa	gccatgaacg	tttacaacga	aaattttcag	caactgaaag	ttcgagttga	4080
	gcatcaactg	gatccgcaaa	actggggcaa	cagcccgaag	ctgcgcttta	ctcgctttat	4140
	caccaacgtc	caggccgcaa	tggacgtaaa	tgacaaagag	ctacttcaca	tccgctccaa	4200
40	cgctcattctg	caccggggcac	gacgtggcaa	tcaggtegat	gtcttctacg	ccgccgggga	4260

agataaatgg aaacgtggcg aaggtggagt acgaaaattg gtccagcgat tcgtcgatta 4320  
 cccagagcgc atacttpaga cgcacaatct gatggctctt ctgtgattca gtgaccattt 4380  
 ttacaaatgg tctactgcaac cgcggtcacc attaatcaaa gggatgtac gtgtatgggc 4440  
 aatcaacaag tcgtttcgat aaccggtgca ggctcaggaa tcggtctcga actggttcgg 4500  
 5tcctttaagt cgcccggtta ttacgtatcc gctctctgac gaaacgagga gcaagaggcg 4560  
 cttctttgca aagagttcaa ggacgcactc gagattgtag tgggcgatgt ccgggaccac 4620  
 gcaacaaatg agaagctgat aaagcaaaca atcgatagat tcggtcatct tgattgtttt 4680  
 attgcaaattg ccggtatctg ggattacatg ctgagcatcg aagagccttg ggagaaaata 4740  
 tcgagcagtt ttgacgaaat attcgacatt aatgtcaaga gctatttcag tggcatcagt 4800  
 10gccgccctgc cggaactgaa aaagactaac ggatcagtgg tgatgaccgc ttcgggtgctg 4860  
 tcccatgctg tcggtggtgg tggttcttgc tacatcgcca gcaagcatgc ggtgctcggg 4920  
 atggttaagg ctttggccta cgaattggcc cccgaagtgc gctgaacgc tgttcgccg 4980  
 gggggcaccg tgacgtctct gtgcggttcc gcgagcgccg gtttcgacaa aatgcacatg 5040  
 aaagacatgc ccggcatcga cgatatgatc aaaggtctca cgcctcttgg gtttcgagcc 5100  
 15aagcccgaag acgtggtggc accctatttg ttgctggctt cgcgaaagca aggaaaattc 5160  
 atcaccggca ccgtgattag cattgatggc ggtatggcgc tcggtcgcaa gtgagcttgt 5220  
 agccgatcag aagttataga cacatttcag gtgacgcccc atgaagacaa aactgtttat 5280  
 caataacgcc tggatcgatt ctagtacca gcagacctc gagcgcatc acccgcagc 5340  
 cagcgatgtg gtgactgaga gcgcaaacgc cacagtgcg gcgcgataa aggcggcgca 5400  
 20agcgcccgag gaggcggtca agacctggaa ggccgttggc ccttcagagc gtcgccgct 5460  
 tctcctaaag gtcgccgatg tcattgaaag taaaacaccc aagttcatcg aagtgatggc 5520  
 catggagggtg ggagcttccg ccttttgggc cggattcaac gtccatgct ctgccaatgt 5580  
 gttccgagag gctgcctcgc tggctaccca aattcagggt gaaaccatcc caacggacaa 5640  
 agccgaaacg ctctcaatga cactacgtca gccggtcggc ccgactctaa gcactgttcc 5700  
 25atggaacggc accgcagtgc ttgcggcacg agccatcgct tatccgctgg tctgtggcaa 5760  
 cactgtgggtg ttcaaaggct ctgaatttag tcccgcgacg catgccctga tcaccagtg 5820  
 cgtgcaggaa gccgggctgc ccgctggcgt gctcaattac ctcaactctt cgcctgaccg 5880  
 ttcgcccgag atcgctgacg cactgatctc tgccaaggag atccgcgcga tcaacttcac 5940  
 gggttccacc cgcgtgggca gcattatcgc gcagaaagcc gcgcaacacc tcaagcgctg 6000  
 30cctgctggag ctcggcgcca agtccccgct tattgttctg gatgatgcag acatcgatgc 6060  
 ggcggctcaag gcagcggtgt tcggtagctt cctgttccaa ggtcagatct gcattgtccac 6120  
 tgagcgcttg atcgttgatg agaagatagc cgacgaattt gtcgcaaaat ttgtcgaaaa 6180  
 aactaagcgc ttgagcgagc gcgaccgtg cgtaactggc gactgcatca tcggcccgat 6240  
 ggtctcgcca aattcgggtg agcggatcaa tggtttgttc aaagacgcga tcgacaaagg 6300  
 35ggcaaaagtt gtttcggcg gcttggccca aggtgcgctc atgccggcca cgatcctgga 6360  
 tcacgtcaaa tctgacatgc ggatttacga tgaggagacc tttggtccca tcaccgtggt 6420  
 aatccgttgt aaaggcgaag cagaggccgt ccgcattgcc aacgacagcg tctatggcct 6480  
 gtcgtcgggc gtatttggcc gcgacatcaa ccgcgctcta cgcgtgggta tgtccatcga 6540  
 atatggttct gtacacatca acggttcgac cgtccagaac gaggcgcagg ctccttacgg 6600  
 40aggcaccaag aacaccggct acgggcgctt cgacggcgt gctgtaatcg acgagttcac 6660

agagatcaag tggctgacca tcgaaccttt cgagcagcaa tatcccttct gataagcact 6720  
 aactcccagg aatcaaaacta tgagtaagca agctgcagtt atcgagctcg gatacatggg 6780  
 tatctcggtc aaggaccttg atgctgggaa atcatttgcc acggatatgc taggtctgca 6840  
 agttcttgat gagggtgaga aggaccgttt ctatctgcgg atggattact ggcacatcg 6900  
 5gatcgtagtc catcacaacg gacaggacga cttggagtac ctaggctggc gtgtagccgg 6960  
 caagccggag ttcgaagctc tgggtcaaaa gcttattgat gccggttaca agatccgcat 7020  
 ctgcgacaaa gttgaggctc aggagcgtat ggtgttgggt ctgatgaaga cagaagatcc 7080  
 gggcggcaac ccgaccgaga tattctgggg ccccggtatc gacatgagca acccgttcca 7140  
 tcccggtcgc cccctgcacg gaaagtgtgt gaccggtgac caaggcttgg gccattgcat 7200  
 10cggttcgcaa accgacgtcg cagaagctca taagttttat agcctgctgg gcttccgtgg 7260  
 ggacgtcgaa taccggattc cgttgcccaa cggcatgact gccgaactgt cgttcatgca 7320  
 ttgcaacgcc cgtgatcact ccattgcttt tgggtgccatg cccgctgcca aacgactcaa 7380  
 tcacttgatg cttgagtaca cccatatgga agacttggga tacacgcacc aacagtttgt 7440  
 aaagaacgaa attgacattg ccttgacgtc tggcattcac gccaacgaca aggcgttgac 7500  
 15gttctatggt gcaacgcctt cgggctggct cattgagccc ggctggcgag gtgccacggc 7560  
 catagatgaa gcggagtatt acgtcggcga catcttcggc catggcgtgg aggccactgg 7620  
 atatggcctg gatgtaaaac tgagctaaag atgcgcgctc gttggcgag gctctagtc 7680  
 agcatcttca tacgcaacca accttgacgg gcgatgagat caaaggacgt taaagcgaag 7740  
 gggaagtgggt tcgggccatg cgcataccga tccatgacat ttgtttcata gtatataggt 7800  
 20agatagggtga atcaagcgtc tagtcaacta gtggacacat ctgttccatg aggcctatcta 7860  
 ctatctattc aaaacaagaa taataaatag gatgaaaata ataatagataa aaagaacgat 7920  
 ttgtcttggt tatcctctat tctgtttggc aagccccaca tgggcccgaag agtcgccttg 7980  
 gacgtaccgt attggtatga ctaatgtagc tttcgatgct agcgcaaaag tatacttaaa 8040  
 tggtcagcgg gtgccaggag gaagcgtga tgcgagcgat aacaacgcgc ttacattcga 8100  
 25cttcggctac gccatcaacg accagtggaa tgtacgtgcg attgtcggta ttccgcctac 8160  
 aactaaagtg acgggcgcag gcacacttcc tgggtatccag ctggggaaaa taacttacgc 8220  
 tccaacagta ttaacgttga actataacct ccccgctttg ggtcccgttc gccctcacat 8280  
 aggtgcggga gtcaattaca cgcggatttt tgaaagtcgg gacgctaata taaaatcggt 8340  
 cgatgccgac cacgcttgggt ccccgcgctc acatgttgggt gccgatattg acgttaaccg 8400  
 30tgggtgggttc gtttagcattg atatccggaa gttatacctg aaaagcgacg catcagggtga 8460  
 cttggggcca caggaggcta aagcacgggt aactcttgac ccattactaa cttcgatcgc 8520  
 gatcggacgc caattctgat gattctgttt aaagtctttt atctatctaa ccgcaaaggg 8580  
 tgtttccatg tcgaataaaa ttatgaaaac gtgcgctctt accgccgaag atatcaacgg 8640  
 cgcttgact ataataccca caccctcgac gcctgatgct tctgattggc gcagcactgc 8700  
 35caccgtggac ttagaagaga ctgcccgcac agttgaagag ctgattgcag ctggtgtcaa 8760  
 cggatttcta agtatgggtg cttttgggtg gtgcgccacg ttgacctggg atgaaaaacg 8820  
 tgattatgtc tcgacgattg tcgagaccat tcgtggtcgc gtgccttatt tctgtggcac 8880  
 gacagcctta aatacccgag aagtcacccg ccagaccgga gagcttatcg atattggcgc 8940  
 caacggcact atgctcgggg tgccgatgtg ggtgaagatg gacctgccta cagcggttca 9000  
 40gttctatcgt gatgttgacg atgcgggtacc agaggctgcc attgcgattt acgccaaccc 9060

```

cgaagcattc aagttcgact tccctcgccc attctgggca gagatgtcca aaattccgca 9120
ggtagtgact gcgaagtatc taggcacggy aatgcttgac ttggacctga gactggcacc 9180
caacatccgc ttccttcccc acgaagatga ctattacgcy gccgcacgca tcaatccgca 9240
gcgcataacc gcgttctggt caagcggggc catgtgcggy ccggctaccg ccatcatggt 9300
5gcgtgacgaa gtggtgcggg ccaagagcac cggtgactgg gccaaggcca aagccatctc 9360
cgatgatatg cgtgcagccg actcgacatt gtttccgcyt gccgactttt cggagttctc 9420
gaagtataat atcgggcttg aaaaggcacg gatggatgcy gctggttggc tcaaggctgg 9480
gccctgccgt ccgccctaca accttggttc agaagactac ctgctggtg cacagaaatc 9540
aggcaaggct tgggcgcgcy tgcacgctaa atacagtaat gaattgaagt agttcacctc 9600
10cgcagacctg agtgacaggg tggcgcagac gctgagggty caggaattaa gtgagctaaa 9660
gcacatttct tgcgccaggc attgccagat cagcaaagtt tgctga 9706

```

```

15      <210> 6
        <211> 2294
        <212> DNA
        <213> Artificial Sequence

20      <220>
        <223> A modified DNA sequence

        <221> unsure
        <222> (186)...(186)

25      <223> n = a or t or g or c

```

```

        <400> 6
agggcagckt ctgtatttgc ggcagcgaaa tgctccctaa attcctcatt taccccatct 60
gaggattgct ttatgacagt aaagtggatt gaagcagtcy ctctttctga catccttgaa 120
30 ggtgacgtcc tcggcgtgac tgtcgagggc aaggagctgg cgctgtatga agttgaaggc 180
gaaatntacy ctaccgacaa cctgtgcacg catggttccg ccgcatgag tgatggttat 240
ctcgagggta gagaaatcga atgccccttg catcaaggtc ggtttgacgt ttgcacaggc 300
aaagccctgt gcgcaccgcy gacacagaac atcaaaacat atccagtcaa gattgagaac 360
ctgcgcgtaa tgattgattt gagctaagaa ttttaacagg aggcaccccg ggccctagag 420
35 cgtaatcacc cccattccat ctttttttagg tgaaaacatg aattacaata ataaatctt 480
ggtaagtga tttggtctga gccaaaagca cctgattcat ggcgatgaag aacttttcca 540
acatgaactg aaaaccattt ttgcgcggaa ctggcttttt ctactcatg atagcctgat 600
tcctgcccc gccgactatg ttaccgcaaa aatggggatt gacgaggtca tcgtctcccg 660
gcagaacgac ggttcgattc gtgcttttct gaacgtttgc cggcatcgtg gcaagacgct 720
40 ggtgagcgtg gaagccggca atgccaaagg ttttgtttgc agctatcacg gctggggcctt 780

```

	cggctccaac ggtgaactgc agagcggtcc atttgaaaaa gatctgtacg gcgagtcgct	840
	caataaaaaa tgtctggggt tgaaagaagt cgctcgcgtg gagagcttcc atggcttcat	900
	ctacgggttc ttcgaccagg agggccctcc tcttatggac tatctgggtg acgctgcttg	960
	gtacctggaa cctatgttca agcattccgg cggtttagaa ctggctggtc ctccaggcaa	1020
5	ggttgatgac aaggccaact ggaaggcacc cgcggaaaac tttgtgggag atgcatacca	1080
	cgtgggttgg acgcacgcgt cttcgcttcg ctcgggggag tctatcttct cgtcgcctgc	1140
	tggcaatgcg gcgctaccac ctgaaggcgc aggcttgcaa atgacctcca aatacggcag	1200
	cggcatgggt gtgttggtgg acggatattc aggtgtgcat agcgcagact tggttccgga	1260
	attgatggca ttcggaggcg caaagcagga aaggctgaac aaagaaattg gcgatgttcg	1320
10	cgctcggatt tatcgagcc acctcaactg caccgttttc ccgaacaaca gcatgtgac	1380
	ctgctcgggt gttttcaaag tatggaacct gatcgacgca aacaccaccg aggtctggac	1440
	ctacgccatt gtcgaaaaag acatgcctga ggatctcaag cgcgcgttgg ccgactctgt	1500
	tcagcgaacg gtcgggcctg ctggcttctg ggaaagcgac gacaatgaca atatggaaac	1560
	agcttcgcaa aacggcaaga aatatcaatc aagagatagt gatctgcttt caaaccttgg	1620
15	tttcgggtgag gacgtatacg gcgacgcgggt ctatccaggc gtcgtcggca aatcggcgat	1680
	cggcgagacc agttatcgtg gtttctaccg ggcttaccga gcacacgtca gcagctccaa	1740
	ctgggctgag ttcgagcatg cctctagtac ttggcatact gaacttacga agactactga	1800
	tcgctaacag acgagtcgac catgatgatc aatattcaag aagacaagct ggtttccgcc	1860
	cacgacgcgc aagagattct tcgtttcttc aattgccacg actctgcttt gcaacaagaa	1920
20	gccactacgc tgctgacctg ggaagcgcat ttgttgaca ttcaggctta ccgtgcttgg	1980
	ttagagcact gcgtggggtc agaggtgcaa tatcaggcca ttacacgca actgcgcgca	2040
	gcttcagagc gtcgttataa gctcaatgaa gccatgaacg tttacaacga aaattttcag	2100
	caactgaaag ttcgagttga gcatcaactg gatccgcaaa actggggcaa cagcccgaag	2160
	ctgcgcttta ctgcctttat caccaacgtc caggccgcaa tggacgtaaa tgacaaagag	2220
25	ctacttcaca tccgctccaa cgtcattctg caccgggcac gacgtggcaa tcaggtcgat	2280
	gtcttctacg ccgc	2294

&lt;210&gt; 7

&lt;211&gt; 4355

30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

35

&lt;400&gt; 7

	atcgataaaa ccatgagggt aataaataat gatgaaggag actatctaga attcgttctg	60
	gagacgggtgc cgggtggcata accgtagacc agcaggctga gcaaaatacc tcgggggtgat	120
	aagccgcgct accgcgtccg ccatatgcct gagtgagtgc gctaaggctg ggccgctcaa	180
40	tggcttcgag aataaagcgg gccagggtgat cttctgtcaa tctcgccgaa tgggtggaag	240

	cacatcaacc	tgaccggcga	ttacgtctgg	cggcagagcc	gcagactgga	ggtcgggaag	300
	ttccggccgt	tacggtggcc	ctgaaaacct	tagtgtagca	ttttttccgt	cttctctatt	360
	cggcctgcgt	tgcgacgttc	gatctgacta	gggcgcagcg	gcaacatcgg	ctgagtacga	420
	tccagcgcct	ggatctgggt	tatttcgtca	accgacaaca	ccagagcggt	atccggtggg	480
5	ttcaggtaaa	gcccaccac	attgactacc	tctcggcaaa	atgagggtcg	ttactaattt	540
	tgaacgtctt	caagcgatgg	ggttttaggt	cgggaagctgc	ccagacctgc	ccaacctacc	600
	acgtcgtgac	ccgagcatat	tccgcccgcg	gcggcaaaaa	acacagaaat	gagcgggtga	660
	cccgatcgcc	tttgatcgat	tctccgcttt	caaaatgggc	gggggctgaa	gtcagccaga	720
	aatacatgac	tacttcagac	gattcgcagt	attcacgctg	gtgataaaca	aattcaacta	780
10	tgctttattg	acaaataaaa	gcacgctcac	catcatcgcg	aatacaaatc	ttataaaaat	840
	taagccggat	ttatggaact	tctcatacag	ccaaacaatc	gcctcattag	ctttagtccc	900
	ggcgccaacc	ttctggaagt	gcttcgcgaa	aacggtgtcg	ctatttccta	cagttgtatg	960
	tctgggcgtt	gcggaacctg	ccgctgtcgg	gttacagatg	gcagtgtaat	tgattcgggg	1020
	gcgggaagcg	ggttaccaaa	cctcgtggac	gagcattatg	tgctcgccctg	tcagtcagta	1080
15	cttactcaca	attgcgcgat	cgaaatccca	gaaaccgacg	aaatcgtcac	ccaccggcg	1140
	agaatcatca	agggcactgt	ggtcgcgctc	gagtcgcccc	ctcacgatat	ccgtgccta	1200
	cgcgtacgcc	tgcctaagcc	cttcgagttc	tcaccgggac	agtacgcgac	attgcagttc	1260
	agtcctgagc	atgcgcgtcc	gtattcaatg	gcaggctctgc	cagatgacca	agaaatggag	1320
	ttccacatac	gcaagggtgcc	gggtgggcgc	gtaacggagt	atgttttcga	gcacgtccgc	1380
20	gaaggtagaa	gcatcaagtt	gagcgggcca	cttggtacgg	cttatttgcg	tcagaaccac	1440
	accgggcccga	tgctctgtgt	gggcgggtggg	accggactag	caccggtgct	gtcgattgtt	1500
	cgcggcgcg	tgaagttggg	tatgacaaac	cccacccctc	tttatttcgg	agtgcgcagt	1560
	cagcaagacc	tctacgacgc	agagcgattg	cacaaactcg	ccgctgatca	ccctcaactg	1620
	accgtacaca	cggtaatcgc	aatggggccc	attaatgaga	gtcagcgagc	cggctctagt	1680
25	accgatgtga	tcgaaaaaga	catcatttcg	ctggctgggt	ggagggccta	cctgtgcggc	1740
	gcaccagcga	tggttgaagc	gctttgcacc	gttaccaagc	atcttggaat	atcacccgaa	1800
	catatttatg	ccgatgcctt	ctatcccggg	ggaatctgaa	tagtcccttt	ccctgcacct	1860
	ctgtccatcg	aggactcatc	aggaggatac	tcaaataagc	gcaaataata	acagtcgcgt	1920
	cagtatttgc	ggcagcgaaa	tggtttccct	ctccctcatt	taccccatct	gaggatagtt	1980
30	ttatgacaga	aaaatggatt	gaagcagtcg	ccctttctga	catcccagaa	ggtgatgtcc	2040
	tcggcgtgac	tgctcagggc	aaggagctgg	cgttgtagca	agtggaaggc	gaaatctacg	2100
	ctaccgacaa	cctgtgcacg	catgggtccg	cccgcagtag	cgatggttat	ctcagggggc	2160
	gagaaatcga	atgccccttg	catcaaggtc	ggtttgacgt	ttgtacaggc	agagccctct	2220
	gcgcccccg	gacagagAAC	atcaaaacat	atgcagtcaa	gattgagAAC	ctgcgcgtaa	2280
35	tgattgattt	gagcggagag	ttttaatagg	aggtagcccg	gaccctagcg	cgtaactacc	2340
	ccaattccat	cttttttagg	tgaaaacatg	aattacaaaa	acaaaatatt	ggtgagtga	2400
	tctgggctga	ccaaaagca	cctgattcat	ggcgatgaag	aacttttcca	gcacgaactg	2460
	agaaccattt	ttgcgcggaa	ctggcttttt	ctcactcatg	acagcctgat	tccatcccc	2520
	ggcgactatg	ttaccgcaaa	aatgggtatt	gacgaggtca	tcgtctctcg	gcagagcgac	2580
40	ggttcgattc	gtgccttcct	gaacgtttgt	cggcaccgtg	gcaagacgct	ggttaacgcg	2640



gaagccggca atgccaaagg tttcgtttgc agctatcacg gctggggctt cggctccaac 2700  
 ggtgaactgc agagcgttcc attcgaaaaa gagctgtacg gcgagtcgct caacaaaaaa 2760  
 tgtctgggggt tgaagaagt cgcctcgctg gagagcttcc atggcttcat ctatggttgc 2820  
 ttcgatcagg aggccctcc tcttatggac tatctgggtg acgctgcttg gtacctagag 2880  
 5 cccatcttca aacattcagg cggtttagaa ctggtcgggc ctccaggcaa ggttgtgatc 2940  
 aaggccaact ggaaggcacc cgcggaaaac tttgtggggg atgcatacca cgtgggttgg 3000  
 acgcacgcgt cttcgcttcg ctcgaggagag tctatcttcg cgtcgctcgc tggcaacgca 3060  
 gtgctgcccc ctgaagggtg aggcttgcaa atgacctcca aatacggcag cggcatgggt 3120  
 gtgttgtggg acggatattc aggcgtgcat agcgcagact tggttccgga gttgatggca 3180  
 10 ttcggcggct ctaagcagga aaggctgaac aaagaaattg gcgatgttcg cgcccggatt 3240  
 tategcagcc acctcaactg caccgttttc ccgaacaaca gcgatgtgac ctgctcgggt 3300  
 gttttcaaag tatggaaccc gatcgacgca aacaccaccg aggtctggac ctacgccatt 3360  
 gtcgaaaaag acatgcccca ggatctcaag cgccgcttgg ccgactcggg tcagcgtacg 3420  
 gttgggcctg ctggcttctg ggaaagcgac gacaatgaca atatggaaac agcgtcgcaa 3480  
 15 aacggcaaga aatatcaatc cagagatagt gatctgcttt caaaccttgg tttcggtaag 3540  
 gacgtatacg gcgacgcggt ctatcctggc gtcgtcggca aatcggcgat cggcgagacc 3600  
 agttatcgtg gtttctaccg ggcttaccag gcacacgtca gcagctccaa ctgggctgag 3660  
 ttcgaggatg cctctagtac ttggcatacc gaactgacga agactactga tcgctaacag 3720  
 acgagtcgac catgatgatc aatattcaag aagacaagct ggtctccgcc cagcagcccc 3780  
 20 aagagtttct tcgtttcttc aattgccacg acgcggcttt gcaacaagaa gccaccacgc 3840  
 tgctgaaccg ggaagcgcat ctgttggaaca ttcaggctta ccgggcttgg ttagagcact 3900  
 gcgtggggtc agaggttcaa tatcagggtca tttcacgcga actgcgcgcc gcttccgagc 3960  
 gccgttataa gctcaatgaa gccatgaacg tttacaacga aaattttcag caactgaaag 4020  
 ttcgaatcga gcatcaactg gatccgcaaa actggagcaa cagcccgaag ctgcgcttta 4080  
 25 ctgcgttcat caccaatgtc caggccgcaa gggacgtaga tgacgaagag ctacttcaca 4140  
 tccgctccaa cgtcattctg caccgggcac gacgtggcaa tcaggctgat gtcttctacg 4200  
 ccgcccggga agacaaatgg aaacgtggcg aagggtggagt gcgaaaattg gtccagcgat 4260  
 tcgtggatta ccagagcgcg atacttcaga cgcacaatct gatggtcttt ctgtgatcca 4320  
 gtgaccactt ttacaaatgg tgactgttac cgcgg 4355

30

&lt;210&gt; 8

&lt;211&gt; 2176

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

35

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

&lt;400&gt; 8

40 gtgatgtcct cggcgtgact gtcgagggtg aggatctggc actgtacgaa gtggaaggcg 60

aaatctacgc taccgacgac ctgtgcacgc atggtgcccgc ccgcatgagc gatggttata 120  
 tcgagggggcg agaaatcgaa tgcccttgc atcaaggteg gtttgacgtt tgtacaggca 180  
 gagccctctg cgcgccctg acagagaaca tcaaaacata tgcagtcaag attgagaacc 240  
 tgcgcgtaat gattgattta agctgagaat ttttaatagg aggcgccccg gaccatagag 300  
 5 cgtaattatc ccattccat ctttttttag gtgaaaacat gaattacaaa aacaaaatct 360  
 tggtaagtga gtctgggctg acccaaaagc acctgattca tggcgatgaa gaacttttcc 420  
 agcacgaact gagaaccatt tktgcgcgga actggctttt tctcactcat gacagcctga 480  
 ttccttcccc cggcgactat gttaccgcaa aaatggggat tgacgaggtc atcgtctctc 540  
 ggcaaagcga cggttcgatt cgtgccttcc tgaacgtttg ccggcaccgc ggcaagacac 600  
 10 tggtgaaagc ggaagccggc aatgctaaag gtttcgtttg cagttatcac ggctggggct 660  
 tcggctccaa cggcgaaactg cagagcgttc cattcgaaaa agagctgtac ggcgagtcgc 720  
 tcaacaaaaa atgtctgggg ttgaaagaag tcgctcgcgt agagagcttc catgggttca 780  
 tctatggttg cttcgatcag gaggccttct ctcttatgga ctatctcgtt gacgtctgct 840  
 ggtacctgga acccatcttc aaacattcag gcggtttaga actggttaggt cctccaggta 900  
 15 aggttgtgat caaggccaac tgggaaggcac ctgcggaaaa ctttgtgggt gatgcatacc 960  
 acgtcgggtg gacgcacgcg tcttcgcttc gcacaggcga atctatcttc tcgtcgtctg 1020  
 ctggcaacgc agtgctgccc cctgaagggtg cgggcttgca aatgacctcc aaatacggca 1080  
 gcggtatggg tgtgttgtgg gacggatatt caggcgtgca tagcgcagac ttggttccgg 1140  
 aattgatggc attcggcggc gctaagcagg aaaggctgaa caaagaaatt ggcgatgttc 1200  
 20 ccgcccggat ttatcgacg caccctcaact gcaccgtttt cccgaacaac agcgtgctga 1260  
 cctgctcggg tgttttcaaa gtatggaacc cgatcgacgc aaacaccacc gaggtctgga 1320  
 cctacgccat tgtcgaaaaa gacatgcccg aggatctcaa gcgccgcttg gccgacgcgg 1380  
 ttcagcgaac ggtcgggcct gctggcttct gggaaagcga cgacaatgac aatatggaaa 1440  
 cagcatcgca aaacggcaaa aaatatcaat ccagagatag tgatctgatt tccaaccttg 1500  
 25 gtttcgggaa ggatgtatac ggcgacgcgg tctatcctgg cgtcgtcgga aaatcggcga 1560  
 tcggcgagac cagttatcgt ggtttctacc gggcttacca ggcacacgtc agcagctcca 1620  
 actgggctga gttcgaggat gcctctagta cttggcatac cgaactgacg aagactactg 1680  
 atcgctaaca gacgagtcga ccatgatgat caatactcaa gaagacaagc tggtatccgc 1740  
 ccatgacgcc gaagagtttc ttcgtttctt caattgccac gactcggctt tgcaacaaga 1800  
 30 agccaccacg ctgctgacct gggaagcgca tctgctggac attcaggctt accggacttg 1860  
 gttagagcac tgcgtggggg cagaggttca atatcaagtc atttcacgcg aactgcgcgc 1920  
 cgcttccgag cgcggttata agctcaatga agccatgaac gtttacaacg aaaatttcca 1980  
 gcaactgaaa gttcgagtcg agcatcaact ggattcacia aactggagca acagcccgaa 2040  
 gctgcgcttt actcgttcca tcaccaatgt ccaggccgca atggacgtaa atgatgaaga 2100  
 35 tctgcttcac gtccgctcca acgtcgttct gcaccgggca cgacgtggca atcaagtcca 2160  
 tgtcttctac gccgcc 2176

&lt;210&gt; 9

&lt;211&gt; 14462

40

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

5

&lt;400&gt; 9

```

gtcgactccg gtggccaaat cgacgggagc acaagcgta cagcatcggg gctatcggcc    60
tcaaccgctg agacgcatct acatacccaa aagcaatggc aagaagcgtc cgctgggaat    120
tccaacgatg cgggatcgtg cgatgcaggc gctgtggaaa ctggcacttg agccggttgc    180
10cgaaacgctg gtagatccca actcctatgg atttcggccg cagcgatcca ctgccgatgc    240
aatcgacat tgcctcaatg cactggcgaa acgtgggtcg gcgcactggg tactggaagg    300
tgacattcga ggctgtttcg acaatatcag tcacgattgg ctgctacca acgtgcccat    360
ggataagggtg gttctgcgca aatggcttcg agcggggtat gtcgatcagg ggccttggtt    420
cgcaacggag gcaggaaccc cgcaaggggg aatcatctct ccggtacttg cgaattggac    480
15gctggatggc ctggaagatg tcgtccatgc aagcgtggct tcgacagcgc gcaagcgtaa    540
gccattcaag atacacgtcg tacgatatgc cgatgacttc atcatcacag gggccacgca    600
agctgtttctg caacatcagg ttctgcctgc aattgaggcg ttctgaaac agcgtggact    660
ggaactctct gatgaagaga ctgagataac gcatatctca caaggcttcg attttctggg    720
ccagaacgta cgcaagtacg ccggcaagct actcatcact ccggtctgca agagtgtgaa    780
20ggcactgctg gataagggtg gggaaatcgc gaacgcaaac aaggcggcga ctcaagctaa    840
cttgatcttg accctgaacc cggtaatccg aggatgggac atgtatcatc gccatgttgt    900
cgccgccaac gtttcgcatg gatcgaccat cagatttggc aagtgttatg gcgctgggca    960
gttcgtcggc acgccatgaa aaatgcctat tgggtaaaac aacgctactt tcgtgtcgta   1020
ggtcaacggc actgggtttt cgccactcag gagaaagcgc gcggtctgag tcaaccgct   1080
25tggtgtgatg ccgcccag tggttcgacg gtgcggcata tcaaaatatg cagtgcagcg   1140
aaccggttcg atccggcatg gacgttttac cttgagcgtc gaagagcgcg tcgtcaggtt   1200
gcccagtctt actcaggctg ctggaaggct tgagccgtat gagggtgac tccaagtac   1260
ggttcttagg ggagacaggg gtggtaacac ccctgtttta cccgactgaa agcgcataac   1320
ggagccgacg cagaatcggg cctggtacat catgtccatg cacagccgcc aacgtggctg   1380
30acgtgacgga agttgccat ctgttgacg gtggcgagaa cgttgtctgt gcttggggtt   1440
acaccggcgt agagaaacgg cctgagcacg acggtcggca agtgatctgg cagatctgcg   1500
cgacgcagca cctacaagca tttgagcaaa cgcagtgcgc cgtacaaggc caggcgcaag   1560
atcgagaagg ccaaagccca ggtacgtgcg aaggtcgaac atccgtttcg ggtgatcaag   1620
cgtcagttcg gttatgtgaa gaccgccttc cgtgtcctgg ccaagaatac ggcggaattg   1680
35accacactgt tcgcgtgtc gaactctgtg gatggtgcgc cggcaattat tgcctgtgc   1740
gggagagggtg cgcccgtagg tgacagaaaa ccagggtttt gcctcggtta tccaataaca   1800
cagcgactga aaatcgggca tttcggcatc cctaagccgc ccatttccga ctgatgagca   1860
acttgttcgg agtttcccta aagcacaagc gcaagccatc agaaaagagt gcctgagtgt   1920
gtccaagat agcccgaaac taaggccgga catggaaaag accaagatca cccatgtgaa   1980
40tgatggactt gtttttggg gccatcggat cattcgcaag cgtagccgat acggagatat   2040

```

gcgtgtgatc acgacgatcc gagacaaagc cagaagattc gcagcctcac tgacggcgct 2100  
 gctatcaggt aaccadagtg aaagtaagat cgacatgggt gaatcagtc accgaaagct 2160  
 taaaggctgg gcggcgttta tcaattcggt gattcaaagc caaagtattc agctatatcg 2220  
 atcgtgtcgt gttttggaaa ctggcccatt ggctggctca aaaataccgc tcccagggtt 2280  
 5gtcttgctg ttcctggccc gccgatgaac accacattct gcgcggtatc tgtgaactcc 2340  
 aggttgacc gctccctatg ggcatcaaca atatctcgtg gccttggtg ggatacgcaa 2400  
 ctaacaaaaa tgcgcgactg gcgcatgcag ctcccatatg ggagacgctg gaagcgtcg 2460  
 aatagaccgt caatcagaag accttctctg ctcccgctga actgaaagcc tcaccgagtg 2520  
 cgaacgtcaa cgggaaaggc gcgcaaagac ttgccttgct cacctcgtca cgagcgtaca 2580  
 10aacgctgtga gctggctgta cgggtgctg gcttagcgtt ggtttttcog ttctgtgagc 2640  
 caaccccggt atatcgtagt gcaggacat gcggttaagg agggctgcct ggtcagcatt 2700  
 ttcagccagc gccggaagtt cctgcacttc ggccttgaac tggcggtcca atgcctggat 2760  
 atggccatag ggaggtacgc tttggcgtaa ggcgttgcat gcagcttggg gctgtcgagg 2820  
 ctgatggtgc ggagcttgag cagcttcatt tgccctgcc aactccaccgc ctgaagggaag 2880  
 15ttcccgcca agtcacgag gaagcgtgc cggaagctgg ccgggggtgac tcttgcgggc 2940  
 aggtggcgaa aagcgagcga gtcgtaactg gcgcgctcaa tctggagct gggaactttg 3000  
 ccggtggcat aaccatagac catcgggctg agcacaatgc ctccgggggtg ataagccgct 3060  
 acccgctccg ccatatgcct gagtggggc gctaaggctg gcccgctcaa tggcttcgag 3120  
 aataaagcgg gccaggtgat cttctgccat ctctcccgaa tgggtggaag cacatcaacc 3180  
 20tgaccagcga ttacgtctgc gcagagccgc agactagagg tcgggaagtt ccggccgtta 3240  
 cggtgcccg gaaaacctca gtgtacgatt ttttcagtct cctctattcg acctgcgttg 3300  
 cgacgttcga tctgactagg gcgcagcggc aacatcggtg gactacgac cagegcctgg 3360  
 atttgggtta tttcggttaac cgacaacacc agagcggtat cctgcgggtt caggcagagc 3420  
 ccaaccacat gactacctct cggcaaaatg agggctgtta ctaattttga acgtcttcaa 3480  
 25gcgatggggg tttaggtcga aagctgcccc ggctgcgga acctaccaca tcgtaaccog 3540  
 agcatattcg gccgcggcg gctaaaaaca cagaatagag cggggtgacc cgatcgctt 3600  
 tgatcgattc tccgctttca aaagcggcag gggctgaagt cagccggaat taccagacc 3660  
 aatcacaaca ttcattgctg tgataaataa attcaactat gctttattga caaataaaag 3720  
 cacactcacc atcatcgca atacaaatct tataaaatta agccggattt atgaaacttc 3780  
 30tcatacagcc aaacaatcgc ctcatagct ttagtcccg gcaccaacct ctggaagtgc 3840  
 ttcgcgagaa cgggtgctgct atttctaca gttgtatgtc tgggcgttgc ggaacctgcc 3900  
 ggtgccgggt tacagatggt agtgtaattg attcggggac ggggaagcggg ttaccacacc 3960  
 tctggagcga gcattatgtg ctgcctgtc ggtcagtact tactaacaat tgcgcgatcg 4020  
 aaatcccaga agccgacgaa atcgtcacc acccgcgag aatcatcaag ggcactgtgg 4080  
 35tcgctattga gtcgccact cagatatcc gtcgcctac cgtacgctc gccaaagcct 4140  
 tcgagttctc acccgacag tacgcgacat tgcagttcag tctgagcat gcgcgtccgt 4200  
 attcaatggc aggtctgcc gatgaccaag aaatggaggt ccacatacgc aaggtgccgg 4260  
 gtgggcgctg aactgagtat gttttcgagc acgtcccgga aggtacaagc atcaaattga 4320  
 gcgggccact tggtagcgt tatttacgtc agaaccacac cgggcccgat ctctgtgtgg 4380  
 40gcggtggaac cgactagca ccggtgctgt cgattattcg cggcgcgctg aagttgggta 4440

tgacaaaacc	catcctcctt	tatttcggag	tgcgcagtca	gcaagacctc	tacgacgcag	4500
agcgattgca	taacctcgcg	gctgatcacc	ctcaactgac	cgtacacacg	gtaatcgcaa	4560
tgggcccgat	taatgagagt	cagcgagccg	gtctagttac	cgatgcgac	gaaaaagaca	4620
tcagttcgct	ggctgggttg	agggcctatc	tgtgcggtgc	accagcgatg	gttgaagcgc	4680
5tttgaccgt	taccaaaccat	cttgggaatat	caccggaaca	tatttatgcc	gatgccttct	4740
atccccgtga	aatctgaatc	gtccctttcc	ctcacctcgg	tccattgagg	actcatcagg	4800
aggatactca	aataggcgca	aataataaca	gccgcgtcac	tatgtgcggc	agcgaaatgg	4860
tttccctctc	cctcatttgc	cccatcggag	gatagtttta	tgacagaaaa	atggattgaa	4920
gcagtcgccc	tttgtgacat	tccagaaggt	gatgtcctcg	gcgtgactgt	cgagggttaag	4980
10gatctggcac	tgtacgaagt	ggaaggcgaa	atctacgcta	ccgacgacct	gtgcacgcac	5040
ggtgcccgcc	gcatgagcga	tggttatctc	gaggggagcg	aaatcgaatg	ccccttgcat	5100
caaggctcgg	ttgacgtttg	tacaagcaga	gccctctgcg	cccccgtagc	agagaacatc	5160
aaaacatatg	cagtcaagat	tgagaacctg	cgcgtaatga	ttgatttaag	ctgagaattt	5220
ttaataggcg	gcgccccgga	ccatagagcg	tgattatccc	cattccatct	ttttttaggt	5280
15gaaaacatga	attacaaaaa	caaaatcttg	gtaagtgaag	ctgggctgac	ccaaaagcac	5340
ctgattcatg	gcggtgaagg	gcttttccag	cacgaactga	gagccgtttt	tgcgcggaac	5400
tggctttttc	tactcatga	cagcctgatt	ccttcccccg	gcgactatgt	taccgcaaaa	5460
atggggattg	acgaggtcat	cgtctctcgg	caaagcgacg	gttcgattcg	tgccttctctg	5520
aacgtttgcc	ggcaccgcgg	caagacactg	gtgaacgcgg	aagccggcaa	tgctaaaggt	5580
20ttcgtttgca	gttatcacgg	ctggggcttc	ggctccaacg	gcgaactgca	gagcggtcca	5640
ttcgaaaaag	agctgtacgg	cgagtcgctc	aacaaaaaat	gtctgggggt	gaaagaagtc	5700
gctcgcgtag	agagcttcca	tgggttcac	tatgcctgca	tcgatcagga	ggccccctct	5760
cttatggact	atctcgggtg	cgctgcttgg	tacctggaac	ccatcttcaa	acattcaggc	5820
ggtttagaac	tggtaggccc	tccaggtaag	gttgatgatca	aggccaactg	gaaggcacct	5880
25gcggaaaact	ttgtgggtga	tgcataccac	gtcggttgga	cgcacgcgtc	ttcgtctctgc	5940
acaggcgagt	ctatcttctc	gtcgtctcgt	ggcaacgcag	tgctgcccgc	tgaagggtgcg	6000
ggcttgcaaa	tgacctccaa	atacggcagc	ggtatgggtg	tgttggtgga	cggatattca	6060
ggcgtgcata	gcgcagactt	ggttccggaa	ttgatggcat	tcggcggcgc	taagcaggaa	6120
aggctgaaca	aagaaattgg	cgatgttcgc	gcccggtatt	atcgacgcca	cctcaactgc	6180
30accgttttcc	cgaacaacag	cgtgctgacc	tgctcgggtg	ttttcaaagt	atggaacccg	6240
atcgacgcaa	acaccaccga	ggtctggacc	tacgccattg	tcgaaaaaga	catgcccagag	6300
gatctcaagc	gccgcttggc	cgacgcggtt	cagcgaacgg	tcgggcctgc	tggcttctgg	6360
gaaagcgacg	acaatgacaa	tatggaaaca	gcatcgcaaa	acggcaaaaa	atatcaatcc	6420
agagatagtg	atctgatttc	caaccttgg	ttcgggaagg	atgtatacgg	cgacgcggtc	6480
35tatcctggcg	tcgtcgaaa	atcggcgatc	ggcgagacca	gttatcgtgg	tttctaccgg	6540
gcttaccagg	cacacgtcag	cagctccaac	tgggctgagt	tcgaggatgc	ctctagtact	6600
tggcataccg	aactgacgaa	gactactgat	cgctaacaga	cgagtcgacc	atgatgatca	6660
atactcaaga	agacaagctg	gtatccgccc	atgacgccga	agagtttctt	cgtttcttca	6720
attgccacga	ctcggctttg	caacaagaag	ccaccacgct	gctgacccgg	gaagcgcac	6780
40tgctggacat	tcaggcttac	cggacttgg	tagagcactg	cgtggggtca	gaggttcaat	6840

	atcaagtc	atc	atc	atc	atc	6900
	ccatgaac	ttgcgac	aat	aactgaa	tcgagtc	6960
	atccacaaa	ctggagca	agcccga	tgcgcttt	tcgcttc	7020
	aggccgca	cgagctaa	gatgaag	tgcttcac	ccgctcca	7080
5	accgggca	acgtggca	caagtcga	tcttctac	cgcccgga	7140
	aacgtggc	aggtggag	cgaaaatt	tgacgcga	cgtggatt	7200
	tacttcga	gcacaatc	atggtctt	tgtgatcc	tgaccact	7260
	gactgtac	gcggtcac	ttaatc	gggaatgt	gtgtatgg	7320
	tcgtttca	aaccggtg	ggctcagg	tcggtctc	actggttc	7380
10	cggccggt	ttgcgtat	gctctcgt	gaaacgag	gcaagagg	7440
	atgaattca	ggacgcac	gagatcgt	tgggcgat	ccgagatc	7500
	agaagctg	caagcaga	atcgctag	tcggtcac	cgattgtt	7560
	ccggtatt	ggattacat	ctgagcat	aagagcct	ggagaaa	7620
	ttgacgaa	attcgacat	aatgtaaa	gctatttc	tggcatcag	7680
15	cggaaact	aaagacga	ggatcggt	tgatgacc	ttcgggtg	7740
	tcggtggt	tggttctt	tacatcgcc	gcaagcat	ggtgttag	7800
	ctttggcc	cgaattgg	cccgaatt	gcgtgaac	tgtctgcc	7860
	tgacgtct	gtgcggtc	gcaagcgc	gtttcgac	aatgcacat	7920
	ccggcatc	cgatatgat	aaaggcct	ctcctctt	gtttgcag	7980
20	acgtggtg	gccctatc	ttgctggt	cgcgaaag	gggaaaat	8040
	ccgtgatt	cattgatgg	ggtatggc	tcggtcgca	gtgagctt	8100
	aggttatag	aacatttt	ggtgacgc	catgaagac	aaactgtt	8160
	ctggateg	tccagtga	agcagacc	cgagcgca	caccccg	8220
	gatgactg	agcgcaaa	ccacggtg	ggacgcga	aaggcggc	8280
25	ggaggcat	aagacctg	aggacgtt	accttcgg	cgtcgccg	8340
	ggtcgcgc	gtcatggaa	gtaaaacac	caagttat	gaagtgat	8400
	gggagctt	gctctttg	ccggattca	cgtccatcg	tctgccaat	8460
	ggctgcct	ctggccact	aaattcagg	cgaaaccat	ccaacggac	8520
	gctctcaat	acactacgt	agccggtc	ccgatcct	agcatcgt	8580
30	caccgcag	ctggcggc	gagccatc	gtatccgct	gtctgtgg	8640
	gttcaaagg	tctgaatt	gtccgcgc	gcatgccct	atcaccag	8700
	agccgggct	cccgtggc	tgctcaact	tctcaact	tcgctgac	8760
	gatcgccg	gcactgat	ccgcgaagg	gatccgcgc	atcaacttc	8820
	ccgcgtggg	agcattat	cgcagaagg	cgcgaacac	ctcaagcgt	8880
35	gctcggcgc	aagtcgcc	ttattgtt	ggatgacgc	aacattgac	8940
	ggcagcgg	ttcggtag	tctgttcca	aggtcagat	tgcatgtcc	9000
	ggtggtgat	gagaagatt	ccgacgaat	tgtcgcca	tttgtcga	9060
	cttgagcgt	ggcgaccct	gcgtaact	cgactgcac	atcgcccaa	9120
	aaattcggg	gagcggat	atggtttgt	caaggacgc	atcgataaa	9180
40	tgttttcgc	ggcatggcc	agggtgcgt	catgccggc	acgatcctg	9240

atctgacatg	cggatctacg	atgaggagac	ctttggctcc	atcacagtgg	tgatccgttg	9300
caaaggggaa	gcagaggcca	tccgcattgc	caacgacagc	gtttatggtc	tgctcgtcggg	9360
cgtgtttggc	cgcgacatca	accgtgctct	gcgagtgggt	atgtcgatcg	aatatggttg	9420
cgtacacatc	aacggctcga	ccgtccagaa	cgaggcgcag	gctccttacg	gaggtaccaa	9480
5gaacaccggc	tacgggcgct	tcgacggccg	tgccgtggtc	gacgagtcca	cagagctcaa	9540
gtggctgacc	attgagccat	tcgagcagca	atatcccttc	tgaagcact	aactccaagg	9600
aatcaaacga	tgagtaagca	agctgcagtt	atcgagctcg	gatacatggg	catctcagtc	9660
aaggatcctg	atgcgtggaa	atcgtttgcc	atgaatatgc	tgggtctgca	agtactcgat	9720
gagggtgaga	aggaccgttt	ctatctgcgg	atggattact	ggcaccatcg	tatcgtagtt	9780
10catcacggcg	gagaggacga	cttggaaatat	ctaggctggc	gtgtagccgg	caagccggag	9840
ttcgaagctt	tgggggcaaaa	gctcattgat	gccggttaca	agatccgtgt	ctgcgacaaa	9900
gttgaggctc	aggagcgtat	gggtgtgggc	ctgatgaaga	cagaagatcc	aggcggcaac	9960
ccgaccgaga	tattctgggg	gccccggatc	gacatgagca	atccgttcca	tccccggccg	10020
cctctgcacg	gcaagtttgt	gaccggtgac	cagggcttag	gtcattgcat	cgttcgccaa	10080
15accgacgtcg	ctgcggccca	taaattctac	agcctgctgg	gcttcctgtg	ggacgtcgaa	10140
taccgcattc	cattgcccaa	cggcatgact	gccgaactgt	cattcatgca	ttgcaacgcc	10200
cgtgatcact	ccattgcgtt	tggtgccatg	cctgctgcca	agcggctcaa	tcacctgatg	10260
cttgagtaca	cccatatgga	agacttgggg	tatacgcatc	aacagttcgt	gaagaacgaa	10320
attgacatcg	ccttgcacgt	tggcattcac	gccaacgaca	aggcgttgac	gttctacggc	10380
20gcaacgcctt	cgggctggct	tatcgagccc	ggctggcgag	gtgctacggc	catagacgaa	10440
gcggagtatt	acgtcggcga	catcttcggc	catggcgctg	aggcacctgg	atatggcctg	10500
gatgtaaaac	tgagttaaaa	gcgattatgc	gtacatcggg	tttctgacgt	tttctgcttt	10560
ctatacagcg	caataacaat	aacaactaag	tgggaatat	aataatgata	aaaaaacgt	10620
ttctttacat	gcctcgcttt	acttgttcac	tcagcacccg	atggggccgaa	gaatcacctt	10680
25ggacataccg	catcggcatg	agtaatat	cttttgatac	aagcgccaaa	gtgtacttgg	10740
gaggccagcg	tgtgccggga	ggaagcgctg	acgcgagcaa	taacaacgcg	ctcacattcg	10800
acttcggcta	tgccatcaac	gaccaatgga	atgcacgttt	gattgtcggg	attccacct	10860
caactaaagt	taaggggcgca	ggcacacttc	cgggcattca	gctgggaaaa	ataacttatg	10920
ctccaacatt	actgacgtta	aactataaacc	tcccagcttt	tggggccggt	cgtcctcaca	10980
30tcgggtgcagg	agtcaattac	acgcgaat	tggaaaagcaa	ggacgcta	ctaaaatcat	11040
tcgatgccga	ccacgcttgg	tcccccgcg	tgcatgttgg	tgccgatatt	gacgtgagtc	11100
gcaactgggt	cgtcagcatt	gatattcgga	agttatacct	gaaaaccgac	gcatcagggt	11160
acttagggcc	acaggaggct	aaagcaaaagg	taactcttga	cccattgata	acatcgatcg	11220
caatcggacg	ccaatttttga	tgcccccttt	taaggctctc	tatctatcta	actgcaaagg	11280
35gtattttttat	gttgaataaaa	attagtaaaa	ccgcgcgtct	taccgctgaa	gatatcaatg	11340
gtgcctggac	tataatgccc	acaccgtcga	cgctgatgc	ttctgattgg	cgcacgacta	11400
acactgtgga	cttagacgag	actgcccgc	tagttgaaga	gctgattgct	gctggtgtca	11460
acggtatttt	gagtatgggt	acctttggtg	agtcgcgcc	gttgacctgg	gaggagaaac	11520
gtgattatgt	ttcgacgggt	gtcgagacca	ttcgcggtcg	tgtaccttat	ttctgcggca	11580
40cgacggccct	gaataaccga	gaagtcattc	ggcagaccgg	agagcttatc	gatattggcg	11640

ctaacggcac catgctagggc gtgcccgatgt ggggttaagat ggacctgccc acagcgggtcc 11700  
 agttctatcg tgatgttgca ggcgcggtac cggaggctgc cattgcgatt tacgccaacc 11760  
 ccgaagcatt caaatcgac ttcctcgc cttttgggc agagatgtct aaaattcctc 11820  
 aggtagtac tgccaagtat ctaggcacg gaatgcttga cttggacctg aaattggcgc 11880  
 5ctaacatccg cttccttcca cagaggacg actattacgc ggccgcacgc atcaatcccg 11940  
 agcgcattac tgcgttctgg tcaagcggg ccatgtgcgg cccggctacc gctatcgct 12000  
 tgcgtgatga agtggagcgg gccaaagagta ccggtgactg gatcaaggcc aaagccatct 12060  
 ccgatgatat gcgtgcagcc gattcgacat tgtttccgcg tggcgacttt tcggagttct 12120  
 cgaagtataa tatcgggctt gaaaaggcac ggaaggacgc ggctggttgg ctcaaggctg 12180  
 10gtccctgccc tctccctac aatcttgttc cagaagatta cctcgttggg gcacagaaat 12240  
 caggcaagtc gtggccgcg ctgcacgcta aatacagtaa agaattaaag tagttcacct 12300  
 ccgcattgct gagcgcgagg ggtgacgtaa acgccgagcg gtgcgggaag taagtgaatt 12360  
 agagttcatt tcttgtgcca ggcactgcta gatcagcaaa gttagctgat ctacgactct 12420  
 cgaaaatttg ggcgaaagct gatcttagga atgcgggata aaggcagtac accgtaacga 12480  
 15tcgggggtg cgttcatgt tgaacgacac gcctattgcg ccgacttctc ttcttcggag 12540  
 tgtttgattg tgattgtcga tttctatttc gattttttga gtccgttctc ttacttggcc 12600  
 aatcagcgtt tgcctaaagct tgcgcaagat catggcctta ccacgtgcta taacgcgac 12660  
 gatttggcgc gggcctaaat agcagtcggt aacggttggc catctaactg cgacttggaa 12720  
 gtaaaatttg actatttgaa agtagatttg caacgggtggg ccagcttcta cggaataccg 12780  
 20ctgggtatttc cagctaacta caacagcaga cggatgaata ctgggcttcta ttactcggag 12840  
 gccaggtgac agcccgctgc ctatgtgaat gtagtattta atgcgatttg gggagaaggc 12900  
 atagcaccag atttggaag cttgcctgccc ctgggtatctg aaaagctagg ctgggacgt 12960  
 agtgccttcg agcgttctc cagcagcaac gccgcaacag agaggtatga cgagcagaca 13020  
 catttcgcca tcgagcga ggtgttcggt gtgccaacga tgtttttggg cgatgaaatg 13080  
 25tggtggggga acgaccgtct gtttatgctc gagagcga tggggcgctt gtgccggaaa 13140  
 aatgccgatt taagtattg atctgatcgt tatttgctcg atgagtcgct ttcaagatca 13200  
 gcgatactg aagtcagcta aatgcgggac tacttcaggc catgcttcag gggcggtcag 13260  
 ggcggactgg cgtgtgcacg gtagttgttg cgtgagcat gtttttaaga tactattttc 13320  
 agtgcctgct gctctctctg ggttgagtgt gagggcgccg atagtctcaa atctcgctt 13380  
 30gtacctgacc agagatacgg gtgtttatga tggcatcgct ggtacagcgc atgacccat 13440  
 gctgcttcac ttcgagtga ttcacgacct cgacgctagc aatctcaacg agccagatga 13500  
 tcgccagcca gggcggaagc tagtcgtgat tcgggttttc tacgttgtac ccgaatgcaa 13560  
 tcttggtctt aacgggtcac gccttaaacy tggaaatgcca aatctgcgat gtcgacattg 13620  
 gccatataa tgttttggt ctctctagc cagaaatcag taagggtgtg atggtgataa 13680  
 35ctgggtcaagc cgggaatcag cggaaatcgt gcgctggcgc cactggcgat caaacagcgg 13740  
 tcaaatata ccattgaact ggcctaagag ggtgtggaca aaataaagta aacttcttgc 13800  
 ctcttttcta tgcagcattg agccatgcct aaaactggac gccaccgta atcgctacgc 13860  
 accattacc tctgctgacc aggcctgctc atacgcagcc ttattccagt caagccgaac 13920  
 tggcgcaaga ttccatgcag aaacgggcat caccgcgcat cccgacacct ttgcaaaagc 13980  
 40gctgagactg gccgggatcg ttcgagtaaa ggagcgtgcc aaaggtagct tccagcctcc 14040



cgagcctcgt aagtettatg gctataccga ggcacaccgg cgtcagttgc cggagcaacg 14100  
 ctatcccagt tgctgactg atgcagaatg gacgttggtc gctgatttat ttgaagtctc 14160  
 gggagggtcgc ggcgtaccgc tcgccactcc cgacgcactc ttttggaaag ctgctgttat 14220  
 gtcgtgcgca cagggtgctc ctggcgaatg ctgcttcgcg agtttcccta ttgggacaat 14280  
 5gtctataaaa cctttcggcg ctggagcgtc agggcaagtt tgagcaaag caccgaccggc 14340  
 tccgtgcccc atggcgagag cgggtggatc gtgatgaacg gccgtcagcc gctgttctgg 14400  
 attcccaatc aacgcgcacg tctccccaag gcggtgaaag cggttacgac gcaggcaaaa 14460  
 ag 14462

10

&lt;210&gt; 10

&lt;211&gt; 12808

&lt;212&gt; DNA

15

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

20

&lt;400&gt; 10

gtcgacgcat gatgcaggac tgggcgcatc gcctcgacct cttcgagcag aaccaggctcg 60  
 aggcggccag catgccgctc accgtgcacg tggaggtgt gcccgcatc ccgagtgcagc 120  
 aaacgcgaag cgcgccctct acgccggttgc cgccttcgcc aagcctgctc gtgacgaagc 180  
 cgggtgacgc catgccgttg gttcttcgcc cgcacatcg gctgccggca gtgccgcccc 240  
 25agcgatcggc cgcgccgctg gtgccttcgg acattcagcg cgagagaatg gatttgttcg 300  
 atgtcttcga agcgcgcac aaccttcccc tcgctgcgtt tgccaagatg gcggggcaaat 360  
 ctgcagggtg gatcagctac gagatccagg cgggcaactt gctggcggtg aacgtgggca 420  
 accgcggcca gcgcgtgccg gactggcacc tcgaccgct caagcacgag ctgatccagt 480  
 ctgtcctgaa gctgaccagg ggtgcggacc cttggcagat ctacatgca ctgctgcagc 540  
 30cgcgctcgat gctgcggggg cgtcggcac tggaggcgt gactgccagc aatctcgaca 600  
 agctcgtcat ggccgtgagc acagcgggtg aggaaagcga atggaccccg ctgcgggtcg 660  
 gtggtcgtgt agtcaaaca cgggcgggta agggatggcc tggacattat tcatattagt 720  
 gatactaata ttcatttatg gtttattgac tattagaata cagtcaatga tcatggaacc 780  
 ataaagcata taaataaaga agctagatct atggaacttc tcgtactacc gaacaatcgc 840  
 35cgcttgectt ttgattccgg tgccaacctt ttggaagtgc tccgtgagca ccgtgtgggt 900  
 atttctaca gctgtatgtc tggacgatgc ggtacttgcg gctgccgagt tatagatggc 960  
 agcgtcatta gttcggcggc gaaaagcggg gactcaaatc gcatcgaaga gcattatgta 1020  
 ctgcctgtc agtcagtgt caccagcaat tgcgcaattg agatcataga ctcagacgac 1080  
 atagtactc acccggcgcg aatcatcaaa ggcatgggtg tcgccgtcga gtcgccact 1140  
 40cacgatattc gccgcacccg cattcgctc gccaaacct tcgagttctc acccggacag 1200

tacgcgatgc	tacagttcag	tcccgaacat	gtgcgtccat	attcaatggc	tggtctgcca	1260
gatgaccaag	aaatggagtt	ccatatccgc	aaagtgccgg	gcgggcgtgt	cacggagtat	1320
atcttcgagc	acgtccgcga	aggtacaagc	attaagttga	gtgggcctct	tggtacggcc	1380
tatctgcgtc	aggctcacac	cgggccgatg	ctgtgtgtgg	gcggcgggac	cggactcgca	1440
5ccgggtgctgt	cgattgttcg	cggcgcgctg	aagtcgggaa	tgacgaaccc	catccacctt	1500
tatttcgggg	tgcgagtcga	gcaagacctt	tacgacgcag	accgattgaa	ccaactcgcg	1560
gctatccacc	ctcaactgac	tgtccatata	gtgatcgcca	cgggcccgat	taatgagggc	1620
cagcggggccg	gcctaattac	cgatttgatt	gaaaaagaca	ttccctcgct	ggctgggtgg	1680
agagcctacc	tgtgcggcgc	accagcgatg	gttgacgctc	tatgcaccgt	cgccaaagat	1740
10cttggaatat	cgcccgagca	tatttatgcc	gacgccttct	atcccagcgg	ggtctgaatt	1800
gccccgaccc	ttcacctctg	tacatcgaga	attcatcagg	aagacactta	aatgagcatc	1860
actaacaaca	gccgcgtctg	aatatattagg	acagctggat	gatctctaac	tccatcatta	1920
ccccatttga	agatagcttt	atgacagaaa	aatggattga	cgcagtcgct	ctttatgaaa	1980
tccctgaagg	tgacgtcctc	ggcgtgacag	togaaggtaa	ggaactagcg	ctgtatgaag	2040
15tgggaaggcga	aatctacgct	accgacaacc	tgtgcacaca	tggtgctgcc	cgcatgagtg	2100
atggctttct	agaaggcaga	gaaattgaat	gtcctttgca	tcaaggtaga	tttgatgttt	2160
gcacaggcag	ggccttgtgc	gcccctgtga	cacagaacat	caaaacatac	ccggtgaaga	2220
ttgagggcca	gcgtgtgatg	attgatttga	gctgagaatt	ttaataggag	gcaccccgga	2280
ccctagagcg	taatcccccc	cattcgatct	cttgagggtga	aaatatgaat	tacaaaaata	2340
20aaaacttgg	gagtgaatct	gggctgaccc	aaaaacacct	gattcatggc	gacgaagaac	2400
ttttccagcg	cgaactggaa	accatttttg	ctcggaactg	gcttttctctg	actcatgaca	2460
gcctgattcc	gtccccctggc	gactatgtta	cggcaaaaat	gggggttgat	gaggttatcg	2520
tctccaggca	gaacgcaggt	tcgattcgctg	cttttctgaa	cgtttgctgt	caccgtggca	2580
agacgctgg	acacgcagaa	gcaggtaatg	ctaaagggtt	cgtttgacgc	tatcacggct	2640
25ggggcttcgg	cgctaacggt	gaactgcaga	gcgtcccgtt	tgaaaaagaa	ctgtatggcg	2700
aggcgctcga	caagaaatgt	atgggattga	agaagtcgc	tcgtgtagag	agcttccatg	2760
gcttcatcta	tggttgcttc	gatgaggaag	ccccttctct	caaagactac	atgggggacg	2820
ctggctggta	cctggagcct	atgtttaagc	attccggagg	gctagaactg	atcggtcctc	2880
caggaaagg	cataatcaag	gctaactgga	aagcgccgc	ggaaaacttt	acgggggatg	2940
30cgtaccacgt	gggttggaag	catgcgtcct	cgcttcgctc	agggcagtcg	gtcttctcgt	3000
cgtagctgg	caacgcagct	ttgccccag	aagggtgcag	tctgcaaatg	acctccaaat	3060
acggcagcgg	catgggtgtg	ttgtgggacg	gatattcagg	cgtgcacagc	gcagacctgg	3120
ttccggaatt	gatggccttc	ggcggtgcta	agcaggaacg	gctgaacaaa	gaaattggcg	3180
aggttcgcgc	acgaatctat	cgcagccacc	tcaactgcac	cgttttcccg	aacaacagtt	3240
35ttctgacctg	ctcgggtgtc	ttcaaggat	ggcaccgat	cgacgcaaat	accactgagg	3300
tatggacctta	cgccatggtc	gaaaaagaca	tgcccagagga	tctcaagcgc	cgcttggtcg	3360
acggggttca	gagaacggtt	gggcctgctg	gcttctggga	aagcgacgac	aacgacaata	3420
tggaaacggt	atcgcaaac	gccaagaaat	atcagtcag	agatggcgat	ctggtttcca	3480
acctgggttt	cggcggggac	gtatacggcg	acgaggttta	tcctggcatc	gtcggcaaat	3540
40cggcgattgg	cgagaccagt	tatcgtggct	tctatcgggc	ttacggcgcg	cacatcagca	3600

	gctctagctg	ggetgaattc	gaggatgtct	ctaaaaattg	gcataccgaa	ctggcaaaga	3660
	ctactgatcg	ctaacdgcg	agagggacca	tgatgattaa	tattcaggaa	gacaagcttg	3720
	tctccgcccc	cgacgccgaa	gagtttcttc	gtttcttcaa	ttccggcgac	gaggctttgc	3780
	aacaagaagc	taccacgttg	ctaaccggg	aagcgcatct	tttagacatt	caggcttacc	3840
	5gcgcctggtt	agagcactgc	gtggactcag	aggtgaaata	tcagattatc	tcacgcgaac	3900
	tgcgctcagc	ttccgagcgc	cgttaccagc	tcaatgaaac	catgaacatt	ttcaacgaga	3960
	attatgaaca	actggaagtt	cgcgtagcgc	atcaactgga	tcgcgaaaac	tggggcaata	4020
	gtccaaaggt	gcgctttact	cgtttcatca	caaatatcca	ggctgcaatg	gacgaaaatg	4080
	aagatttget	tcacattcgc	tccaacctaa	ttgttcacgc	agcacgacgc	ggcaatcaag	4140
10	tcgatgtctt	ctatgccact	cgggaggata	aatggaagcg	cggcggaagat	ggagcgcgta	4200
	agttggtcca	acgattgatt	gattatccag	agcgcacatt	ccagacgcac	aatgtgatga	4260
	tctttatgtg	acccaataat	cgcctttaca	aatggtgact	gctacaagcg	gtccatttgt	4320
	tcaaaaggaa	atttatgtgt	atgagcaatc	aacaagtcgt	ttcgataacc	ggtgctggct	4380
	caggaattgg	tctcgaactg	gttcgatcct	ttaagtccgc	cggttattgc	gtatccgctc	4440
15	tcgtacaaaa	cgaggagcaa	aaggcgagcc	tttgcaatga	gttcaaggac	gcactcgaga	4500
	tcgtcgtggg	cgatgtccgg	gaccacgcaa	caaagtagaa	gctgataaag	caaacaaccg	4560
	atagattcgg	ccatctcgat	tgtttcattg	caaatgccgg	tatttgggat	tacatgcttg	4620
	gcacgaaga	gccttgggag	aaaatatcga	gcagttttga	tgagatattc	aacatcaatg	4680
	tcaagagcta	tttcagcggg	atcagggccg	ccctgcagga	actgaaaaag	actagcggat	4740
20	cagtggatgat	gaccgcttca	gtgtcgtccc	atgcggtcgg	tgctgggtgg	tcttgetaca	4800
	tcgccagcaa	gcacgcgggc	ctgggcatga	tgaaagcttt	ggcttacgaa	ttggctcccc	4860
	acattcgcgt	caacgccgta	gcaccggggc	gcactgtgac	gcctctgagc	ggtcccgcga	4920
	gcgccggcgt	cgacaaaact	cacatggaaa	acatgcccgg	tatcgaggac	atgatcaagg	4980
	gtctaacgcc	tcttggaatt	tcagccaagg	ccgaagacgt	agtggcaccc	tatttgttgt	5040
25	tgccgctcgcg	agatcaaggg	aaattcatta	ccgggactgt	cattaatata	gatggaggga	5100
	tggcgtcgcg	tcgcaagtag	gtttgtcgcc	tatcttgaaa	taataactaa	atttctggta	5160
	aaaccgcatg	aatacaaaat	tgtttatcaa	caatgtctgg	atcaattcca	gtgaccaaca	5220
	gaccttcgag	cgaaagcacc	ccgtcagtgg	tgaggtgatg	acggagtgtg	caaactccac	5280
	ggtgatggat	gcgttaaagg	ccgcgcaagc	tgcccaagag	gctttccaga	cctggaagac	5340
30	tgttggaact	tcggagcgtc	gccgccttct	gctgagggtc	gctgagggtta	tggaaagtaa	5400
	aacacccgag	tttatcgaag	tgatggccaa	ggaggtggga	gcctccgctc	tttgggccgg	5460
	cttcaatgtc	cagatgtcag	ccaatgtgtt	ccgtgaagcg	gcacgctggg	ctacacaaat	5520
	tcagggggaa	actattccga	cagacaagtc	tgacacgctc	tcaatgacgc	tacgtcagcc	5580
	ggtcggtcgc	atcctgagca	tcgtgccgtg	gaacggcacc	gcagtgtctg	cggcacgagc	5640
35	catcgcttat	ccgctgggtc	gcggcaacgc	ggtggtatct	aaaggttctg	agtttagtcc	5700
	cgcgacgcat	gccctgatca	cccagtgcgt	gcagggaagcc	gggctgcctg	ctggcgtgct	5760
	caactatctc	aactcttcgc	ctgaccgttc	gcccagagac	gccgacgcac	tgatctcagc	5820
	caaggagatc	cgacgcatca	acttcacggg	ctccaccgcg	gtgggcagta	ttatcgcgca	5880
	gaaggccgcg	caacacctca	agcgtgcctc	gctggagctc	ggtggcaagt	ccccacttat	5940
40	gttcttgat	gatgcagaca	tcgatgcggc	ggtcaaggca	gcgggtgttcg	gtagcttctc	6000

gttccaaggt	cagatctgca	tgtccactga	gcgcttgatc	gttgatgaga	agatagccga	6060
cgaatttgtc	gcaaaatttg	tcgaaaaaac	taagcgcttg	agcgcaggcg	accggtgctg	6120
aactggcgac	tgcatcatcg	gcccgatggg	ctcgccaaat	tcgggtgagc	ggatcaatgg	6180
tttgttcaaa	gacgcgatcg	acaaaggggc	aaaagttgtt	tgcggcggct	tggcccaagg	6240
5tgcgctcatg	ceggccacga	tcctggatca	cgtcaaactc	gacatgcgga	tttacgatga	6300
ggagaccttt	ggccccatca	ccgtggtaat	ccgttgtaaa	ggcgaagcag	aggccgtccg	6360
cattgccaac	gacagcgtct	atggcctgtc	gtcgggcgta	tttgggcgcg	acatcaaccg	6420
cgctctacgc	gtgggtatgt	ccatcgata	tggttctgta	cacatcaacg	gttcgaccgt	6480
ccagaacgag	gcgcaggctc	cttacggagg	caccaagaac	accggctacg	ggcgcttcga	6540
10cgggcgtgct	gtaatcgacg	agttcacaga	gatcaagtgg	ctgaccatcg	aacctttcga	6600
gcagcaatat	cccttctgat	aagcactaac	tcccaggaat	caaactatga	gtaagcaagc	6660
tgcagttatc	gagctcggat	acatgggtat	ctcgggtcaag	gacctgatg	cgtggaaatc	6720
atttgccacg	gatatgctag	gtctgcaagt	tcttgatgag	ggtgagaagg	accgtttcta	6780
tctgcggatg	gattactggc	atcatcggat	cgtagtccat	cacaacggac	aggacgactt	6840
15ggagtagccta	ggctggcggtg	tagccggcaa	gccggagtgc	gaagctctgg	gtcaaaaagct	6900
tattgatgcc	ggttacaaga	tccgcatctg	cgacaaagtt	gaggctcagg	agcgtatggg	6960
gttgggtctg	atgaagacag	aagatccggg	cggcaaccgg	accgagatat	tctggggccc	7020
ccggatcgac	atgagcaacc	cgttccatcc	cggctcgccc	ctgcacggaa	agtttgtgac	7080
cggtagccaa	ggcttggggc	attgcatcgt	tcgccaaaac	gacgtcgag	aagctcataa	7140
20gttttatagc	ctgctgggct	tccgtgggga	cgtcgaaatac	cggattccgt	tgcccaacgg	7200
catgactgcc	gaactgtcgt	tcatgcattg	caacgcccgt	gatcactcca	ttgcttttgg	7260
tgccatgccc	gctgccaaac	gactcaatca	cttgatgctt	gagtacaccc	atatggaaga	7320
cttgggatac	acgcaccaac	agtttgtaaa	gaacgaaatt	gacattgcct	tgcagcttgg	7380
cattcacgcc	aacgacaagg	cgttgacggt	ctatggtgca	acgccttcgg	gctgggtcat	7440
25tgagcccggc	tggcgagggtg	ccacggccat	agatgaagcg	gagtattacg	tcggcgacat	7500
cttcggccat	ggcgtggagg	ccactggata	tggcctggat	gtaaaactga	gctaaagatg	7560
cgcgctcggt	gggcgaggct	ctagtccagc	atcttcatac	gcaaccaacc	ttgcaggggc	7620
atgagatcaa	aggacgttaa	agcgaagggg	aagtgggtcg	ggccatgcgc	ataccgatcc	7680
atgacatttg	tttcatagta	tataggtaga	taggtgaatc	aagcgcttag	tcaactagtg	7740
30gacacatctg	ttccatgagg	ctatctacta	tctattcaaa	acaagaatga	taaataggat	7800
gaaaataata	atgataaaaa	gaacgatttg	tcttggtgat	cctctattct	gtttggcaag	7860
ccccacatgg	gccgaagagt	cgccttggac	gtaccgtatt	ggtagacta	atgtagcttt	7920
cgatgctagc	gcaaaagtat	acttaaatgg	tcagcgggtg	ccaggaggaa	gcgctgatgc	7980
gagcgataac	aacgcgctta	cattcgactt	cggctacgcc	atcaacgacc	agtggaaatg	8040
35acgtgcgatt	gtcgggtattc	cgctacaac	taaagtgcag	ggcgcaggca	cacttcctgg	8100
tatccagctg	gggaaaataa	cttacgctcc	aacagtatta	acgttgaact	ataacctccc	8160
cgttttgggt	cccgttcgcc	ctcacatagg	tgcgggagtc	aattacacgc	ggatttttga	8220
aagtggggac	gctaatactaa	aatcgttcga	tgcgcaccac	gcttgggtccc	ccgcgctaca	8280
tgttggtgcc	gatattgacg	ttaacctggg	ttggttcggt	agcattgata	tccggaagtt	8340
40atacctgaaa	accgacgcat	cagggtactt	ggggccacag	gaggctaaag	cacgggtaac	8400

tcttgaccca	ttactaactt	cgatcgcgat	cggacgcca	ttctgatgat	tctgtttaa	8460
gttctttatc	tatctaacg	caaaggggt	ttccatgtcg	aataaaatta	tgaaaacg	8520
gcgtcttacc	gccgaagata	tcaacggcg	ctggactata	atgcccacac	cctcgacg	8580
tgatgcttct	gattggcgca	gcactgccac	cgtggactta	gaagagactg	cccgcata	8640
5tgaagagctg	attgcagctg	gtgtcaacgg	tattctaagt	atgggtactt	ttggtgag	8700
cggcacgttg	acctgggatg	aaaaacgtg	ttatgtctcg	acgattgtcg	agaccatt	8760
tggtcgcg	ccttatttct	gtggcacgac	agccttaaat	acccgagaag	tcatccgcca	8820
gacccgagag	cttatcgata	ttggcgccaa	cggcactatg	ctcgggggtg	cgatgtgg	8880
gaagatggac	ctgcctacag	cggttcagtt	ctatcgtgat	gttcgagatg	cgggtacc	8940
10ggctgccatt	gcgatttacg	ccaacccgga	agcattcaag	ttcgacttcc	ctcgccatt	9000
ctgggcagag	atgtccaaaa	ttccgcaggt	agtgaactcg	aagtatctag	gcacggga	9060
gcttgacttg	gacctgagac	tggcacccaa	catccgcttc	cttccccacg	aagatgacta	9120
ttacgcggcc	gcacgcatca	atcccgagcg	cataaccgcg	ttctggtcaa	gcggggccat	9180
gtgcggcccg	gctaccgcca	tcatgttgcg	tgacgaagtg	gtgcggggcca	agagcaccg	9240
15tgactggggc	aaggccaaag	ccatctccga	tgatatgcgt	gcagccgact	cgacattg	9300
tccgcgtggc	gacttttcg	agtctcga	gtataatc	gggcttgaaa	aggcacggat	9360
ggatgcggct	ggttggctca	aggctggg	ctgcgctcg	cctacaacc	ttgttcaga	9420
agactacctc	gctgggtgc	agaaatcagg	caaggcttg	gccgcgctg	acgctaaata	9480
cagtaatgaa	ttgaagtagt	tcacctccg	agacctgagt	gacaggggtg	cgcagacg	9540
20gaggggtgc	gaattaagt	agctaaagca	catttcttgc	gccaggcatt	gccagatc	9600
caaagtttgc	tgatctggca	gtttcaaaaa	tttgggcgaa	agctgatatc	aggaatac	9660
gataaaggca	gtgcaccata	acgacgggg	cgtgccattc	gtgatgaacg	attttgctat	9720
tgtgcgact	tctgttcttg	gagtgtttga	ttgtgattgt	cgatttttat	ttcgattttt	9780
tgagtccgtt	ctcttacttg	gccaacctc	gtttgtcaaa	gcttgcgcaa	gactatgg	9840
25tttccattcg	ttattacgca	atcgatttgg	cgcgagttaa	aatagccatc	ggaaaacgtg	9900
gtccatctaa	tcgcgacctg	atagtcaagc	tggactat	gaaagtagat	ttgcaacg	9960
gggcgagct	ttacgaaata	ccgttggtat	tcccagctaa	ctacaacagc	cgacggatga	10020
atactgggct	ttattactcg	ggagccatgg	cacagactgg	tgctatgtg	aatgtagtat	10080
ttaatgcggt	ttggggagat	ggcatagctc	cagatttgg	aagcttgcc	gctctggtat	10140
30ctgaaaaact	aggctgggat	cgtagcgct	tcgaggactt	tatcagcagc	gatgccgcaa	10200
cagagaggta	tgacgagcag	acacatgccg	cgatcgaacg	caaagtgttc	ggtgtgccaa	10260
cgatgttttt	gggcgatgaa	atgtgggtgg	gaaacgaccg	tctatttatg	ctcgagaacg	10320
cagtgggagg	tgcgctgtga	aatggagaat	agtcgctacg	gagcgcttgt	gccggctaaa	10380
tgccgatata	agtggttgac	ctgatcggtta	tttgctcgat	acagcgcttt	caaaatcagc	10440
35ggctactgaa	gtcagataaa	aatgcgggac	tacttcaggc	atcctgtgcg	acacaaag	10500
ttacctgtaa	ttgtccacct	attccgagtt	tggaatggta	gctgactcgc	tatgcgacca	10560
gcgatagcct	aacaagacat	gcatactgg	taacggtggg	gtgtgaagct	cctgcaacaa	10620
tgtagcccct	tgatgtgtgt	atttgctcg	aggtgaagca	cagatgctcg	gagccgtacc	10680
ggcttgtggc	gctaggtcgg	caagtatgag	caacgtaagt	gggggttggg	gcgcaatggg	10740
40aaccaaaaac	caacgcaagc	cttaccagcg	tcgttcgggtg	ccttcctccc	atgcctccgc	10800

ctcgataaag cagctgcgca tatecggttc ctggctgata tcggttagta ggatcatgcaa 10860  
 ggtcttgtcc agcgectcgt cgctccgata cggaatggtc agctcgtaat ggccgggtctc 10920  
 cgaccgcttc atgccgtagg gctccaggca gtagcgctcg atgttctccg tggcccgtctt 10980  
 ccgaccgcgc atgaacttgc tgttgttcac caccgccagg cgcagggtga cgggtggccac 11040  
 5ccgctcgacg gttgactctg ccggtgacgc gatattgcgc ttttgacctc gcgccagggc 11100  
 gctcttcttg tacgtcccga tctcgacgcc acgggtggcg aggtagctgt acagggtgct 11160  
 cttggagatg tgcaacttct cgccgatggc gctgacgctc aggcggccct cgcggtacag 11220  
 ggtctccgcc gccatggcgg tggcctcggc cttggctggc aggcctcttg gacggcgacc 11280  
 gatccggcct cgagtccgtg ccgccgacag gcccgctga gtccgctcgc ggatcagctc 11340  
 10gcgctcgaa tgaacaggtt gaacaccagg cgatcttggg cgtgggtgct gtcaatgggg 11400  
 tcgttcaggc tctgcaagcc gactttgcgt gcagccagct agccgaccaa ctcaaccagg 11460  
 tgcttgagcg agcgaccgag gcgatccagc ttccagatca ccacggcatc gcccgctcga 11520  
 acatgggcta gcaacttgc caactccggc cgcgcgcttt ttgcgcgct ggcgatgtct 11580  
 tgatagatgc gttcgcaccc ggctgtttc agggcatcga cctggaggtc ggcgttctaa 11640  
 15tcccgagtgc tcaccgcgt ataaccgac ttcataaaaa gtaccgttta cttgactgca 11700  
 ttagtaatag ttgaactttg attaaagctta ccagttattt gaacctagc gcagggtgaa 11760  
 ggcctcagcc gcccacctc tactcagctt gatgaaccga ggggcagtag ttcattcaatc 11820  
 cggctgttag ggaaactctg aagaagactt cctgattttg gcaaaatgcc cggatttcac 11880  
 ccgcgagtt ttccaatgaa gcagatgacc ttctcgcagc ccgagtatgc cggtaagcgc 11940  
 20aaacagaccc gcaaagagct gttcctgata gagatggatc ggggtggggt gtggaagggt 12000  
 ttgattgctt tgatcgatcc gcattacccc aagggtgagg gtggcccgcc agcctatcca 12060  
 ctgacggcaa tgctgcggtt tcactcgatg cataagtggg tcggctacag cgaccggcg 12120  
 atggaggaag cgtgttacga gacgaccatc ctgcgccagt ttgccgggtt gagcctggag 12180  
 cgcattccgg actaaaccac catccccaac ttccgcggc tgctggagaa acacgaacta 12240  
 25cctaccggca tcatgtttgc tgaattttcg cctttggctt tgagtgcggc attcaaccga 12300  
 tacagggtgt gcattgacag ccaaagaaat caaccttggc ggctagaaca cttgtcgatg 12360  
 aagcgaacaa tccgcaaacy ttaaggctac caggaaaaac ccaggatgcg gtagccgata 12420  
 gcccgcgcg gcattgtct ccttcgctc aagagtgtca aaggatattg gactaatgca 12480  
 gcatgcctcg accacatggc cgtgggtcgt tcatctgttg gccatacgtc tatcgaatcg 12540  
 30gacgcgggcg gcgacgacct cgggttgttg gcgttcggcc cagtcgatga gctcagtcag 12600  
 ggacggcatc agcgaagtcc cgagcggggt gagcccgtag cgcacagatg ggggtgtcga 12660  
 gggcggtacc tcgcgccaga tgagtccgtc gcgttcgagg tgacgcagcg tctcggtcag 12720  
 catccggcgc gaaatgtcgg gcacggcgcg ggcagcgca ttgaatcgtt gcggaccctg 12780  
 cgacaagggt accaggatca gcgtcgac 12808

35

&lt;210&gt; 11

&lt;211&gt; 15024

40

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

5

&lt;400&gt; 11

atcgtcctta agcgccgcac gcagccgttt aagtgaattg ctgacggcag gctgctcag	60
ccccagtttt tcgcccggcg tcgatacgtt ccggtcgagc agtagctggt tgaagaccac	120
cagcagattc aagtcgatgt cgcgcagatc catgatgcct caccattatt catgctggtg	180
10attttaacta tcagacttga tctatagcgc tataccgatc gacgcgccag aatcgcagcc	240
attcggagac aactgaaaaa agagcttgca tggaaactgg agtagaacce ctcaatttgc	300
atctgaacgc ggagaccggc agcaccctgc ttgacgtgct caggtccaac gaggtcccca	360
ttttttatag ctgcatgtcg ggcgctgctg gcacttgccg ttgcccgtgtg attgccggcc	420
atcttcgcga taacggcccc gagacagggc gcccgcaggc aggaaaggga acctacgttc	480
15tggcctgtca ggcggttctg accgaagact gcacgatcga gattcctgaa tctgacgaga	540
tcgtggttca cccggcgcgc atcgtcaagg ggacggtcac agcgatagac gaagccaccc	600
atgacatccg gcgcctgccc atcaaactgg ccaaaccgct tgagttcagc cctggccagt	660
acgcaacggg gcagttcacg cccgaatgcg tccgcccata ttcgatggcc gggctgccta	720
gcgatgcgga aatggagttt cagattccgc cggttccggg cgggcatgtc agcaactacg	780
20ttttcaatga actgtccgta ggcgcttcgg tgcggatcag cgccccctc ggaacggctt	840
atctgcggcg cacgcacacc ggccccatgc tttgtgtggg ggggtggaaca ggtctggcgc	900
ccgtcctttc gatcgttcga ggcgcactgg aaagcgggat gagcaacccc atccatctgt	960
acttcgggtg gcggagcgag caggacatct atgacgagga acgccttcac gcattggctg	1020
caaggtttcc gaatctcaag gtgaatgtcg ttgttgcaac aggcctgcc ggccctggtc	1080
25gtcgatccg cctggtcacc gatctgatcg gccgtgactt gcccaatttg gcgggatggc	1140
gcgcctacct gtgtggcgct ccggccatgg tcgaggccct gaacctgctc gttgctcgcc	1200
taggcatagt acccgggcac atccatgccg atgcgttcta tcccagcggc gtctgagcga	1260
aggcaccatg cgaacccaat tcaacccaag gataccaagc catgagtga ccccaacgat	1320
taaaaccctg gtttcccca gatccgaaat ggccgggcca aggtagcagc cgcgttcctt	1380
30tctgggccta caccgcgaa gacctgtaca agcgcgaatt ggagcgctg ttctatgcaa	1440
accactggtg ctatgtaggc ctggaagccg agattccgaa tccaggcgac ttcaagcgaa	1500
cggatgatcg tgagcgctcg gtcacatggt tgcgtgatcc ggatggcggc atcaacgtgg	1560
tggagaacgt ctgcgcccac cgtggcatgc gcttttgccg cgagcgccac ggcaacgcca	1620
aggacttctt ctgcccctac caccagtgga actacagcct caagggtgac ctgcagggcg	1680
35tgccttccg ccgtggcgctc aagcaggacg gcaagggtcaa cggcggcatg cccaaggact	1740
tcaaactcga agaacacggc ctgaccaagc tcaagggtgc cggccgaggc ggtgcagtgt	1800
ttgcctcttt tgaccacgat gtcgagccct tcgaggagtt cctgggcccc accatcctgc	1860
attacttcga ccgcgtcttc aacggccgca agctcaagat cctgggctac cgccgcccgc	1920
gcaccccggt caactggaag ctgatgcagg agaacaatcaa ggacccctac caccggggcc	1980
40tgctgcacac ctggttctcg accttcgggc tctggcgcgc cgacaacaag tcggaactga	2040

agatggacgc	caagttccgc	cacgcgcgcaa	tgatctccac	gcgcgggtcag	ggcgggaaga	2100
acgaggaggt	cgtgtccggc	gtggacagct	tcaaggaaca	gatgaagggtg	aacgaccgc	2160
gcctgctcga	catcgtgcc	gagccctggt	ggggcgggtcc	gactgcgggtg	atgaccacga	2220
tcttccccag	cgtgatcctc	cagcagcagg	tcaacagcgt	atcgaccgc	cacatccagc	2280
5ccaacgggtca	cggctccttc	gatttctgtct	ggaccactt	cggcttcgag	gacgacaacg	2340
aggagtggac	ccagcgcgc	ctgatccagg	ccaacctgtt	cgggcggcg	ggcttcgtgt	2400
cggccgatga	cggcgagggtg	atcgagtgggt	cgcaggaagg	ctttgagcaa	aaaccgacgc	2460
accgcaccgt	gatcgagatg	ggcggtcacg	aaatcggcga	cacggaccac	atggtcaccg	2520
agacgtgat	ccgcggcatg	tacgactact	ggcgcaagggt	gatgggggaa	taaacaatggt	2580
10cgacttcaaa	acctatttgc	aactgctgaa	cctgtacagc	gactacgcca	tggtgtgcga	2640
ctccgccaat	tgggagaagt	ggcctgattt	cttcacgag	accggcacct	accgcctgca	2700
gccgcgcgag	aactttgagc	agggcctgcc	gctgtgcctg	ctggcgctgg	agagcaaggc	2760
aatgatccgc	gaccgggtgt	acggcgtaaa	ggaaaccatg	taccacgacc	cctactacca	2820
gcgccacatc	gtgggcacgc	cgcgcgtgct	gtcagtggag	cgtgacgcgg	atggcgagcg	2880
15catcacgcc	gaagccagct	atgccgtgat	tcgcaccaag	tacgacggcg	attccacgat	2940
tttcaacgcc	ggctattacc	gagacgtgat	cgtgcgcacg	cccaggggcc	tcaagctgaa	3000
gtcgcgcctg	tgcgtttacg	acagcgaaat	gatccccaac	tctgtgatct	atccaatctg	3060
aggcgacgca	catgactcag	aactggattg	atgcagcttg	tcttgacgac	atccctgaag	3120
gcgatgtggt	cggcgtaaaa	gttaacggca	aggaaattgc	gctctacgag	gtcgaagggg	3180
20agatttatgc	caccgataac	ctgtgcacgc	acggcgctgc	gcgcgatgagc	gatggctttc	3240
tcgaaggccg	ggaaattgaa	tgcctcttgc	accaaggctc	atttgatgtc	tgacacggca	3300
aagccttggt	cacccccctg	acaaaggaca	tcaaaacctc	ccccgtcaaa	attgaaaaca	3360
tgcgctgat	gctcaaaatg	gagtaagact	gctttcaatc	aggaggtgta	attcggaaca	3420
atcaacaaa	ccatccaatc	caccatttag	ccaaaaacg	gagacattaa	tcatgattta	3480
25tgaaaatttg	gtgagtgaag	cagggtgac	gcaaaagcac	ctgattcatg	gcgacaaaga	3540
acttttccag	cacgaattga	agaccatctt	cgcgcgaaac	tggctttttc	tgaccatga	3600
cagcttgatt	ccctccccgc	gcgactatgt	cacagccaaa	atgggtgtcg	atgaagtcac	3660
cgtctccgc	cagaacgatg	gctcgggtgc	agcctttttg	aatgtttgcc	gtcaccgggg	3720
caagacacta	gttcacgctg	aagccggaaa	tcgaaaaggc	tttgtgtgca	gttaccacgg	3780
30ctggggcttt	ggctccaacg	gcgaactgca	aagcgttccc	tttgaaaaag	agttgtacgg	3840
cgatacgatc	aaaaagaagt	gcctgggctt	gaaagaagt	ccccgcacgc	aaagctttca	3900
tggcttcatt	tacggttggt	ttgatgcaga	ggccccacg	cttgctgatt	atctgggtga	3960
tgacgcctgg	tacttagaac	ccatcttcaa	gcactctggt	ggcctggaac	ttgtaggccc	4020
ccccggcaaa	gtggtgatca	aggccaactg	gaaggctcct	gcggaaaact	ttgtgggtga	4080
35cgcgatccac	gttggttgga	cgcacgcac	gtctttgcgc	tcaggtcagt	cgatatttac	4140
ccctcttgcg	ggcaacgcta	tgcttccacc	cgaaggcgcg	ggcttataaa	tgaccagcaa	4200
gtatggcagt	ggaatgggcg	tattgtggga	cggctactcc	gggtgtccaca	gtgctgacct	4260
ggttcccgaa	atgatggcat	tcggcggcgc	aaaacaggaa	aaactcgcca	aggaaatcgg	4320
cgatgtccgg	gcacggattt	accgcagcca	tctaaactgc	acggttttcc	cgaacaacag	4380
40cattttgacc	tgctccggtg	tcttcaaggt	ctggaaccgc	atcgatgaaa	acacgaccga	4440



ggtttggacg tatgccatcg tagaaaaaga catgectgag gacttaaagc gtcgcttggc 4500  
 tgacgcggtt cagcgdactg tccgaccagc aggattctgg gaaagcgacg acaacgacaa 4560  
 catggagacg gagtcgcaaa atgccaagaa ataccaatcc agcaacagtg atctgattgc 4620  
 caatttgggt ttcggcaagg acgtctacgg cgacgaatgc tatccgggcg tcgttgccaa 4680  
 5atcggcaatc ggcgaaacca gctatcgagg attctaccgt gcctaccagg ctcacatcag 4740  
 cagctccaat tgggcccagc tcgaaaacac ctcccgaat tggcacaccg aactcaccaa 4800  
 gacgactgat cgctaatacca ggagccaacc atgatgatca ataccagga agacaagctg 4860  
 gtctccgcgc acgacgccga agaatttcat cgtttcttcg tcgggcacga cagcgatctg 4920  
 cagcaagaag tcaccacact cctgaccoga gaagcacatc tgttggacat tcaggcctac 4980  
 10aacgcctggc ttgaacactg cgttgccccc gagatcaaat accaagtgat ctgcgagaa 5040  
 tttcggcca cttccgagcg tcgataccaa ctgaatgatg cgggtgaacat ctacaacgag 5100  
 aactatcagc acctgaaagt tcgagttgag catcagatgg acccgagaa ctgggccaac 5160  
 agcccgaaga tccgtttcac ccgcttcgtc accaatgtca cagcggccaa ggacaagatc 5220  
 gtaccggatc tgctgcatgt gcgttccaac ctattctcc accgcgccag acgcggaat 5280  
 15caagttgacg tcttctatgc aacgcgggaa gacaaatgga aacgcacga aggtgggtggc 5340  
 atccaactgg tggaaactct tgtggactac ccggagcgca ttctccagac ccacaatctg 5400  
 atgaccttc tgtgaacct ggggatgctt gccttgatgg cgggtcatct tgattgtttt 5460  
 aaacagaaat ttattgccat gaacatacag caagttattg ccattactgg cgccgggttc 5520  
 ggcacggggt tggaaactgg tcgatccttc aaagcagctg gctattgctg gtccgcactt 5580  
 20gttcgcaacg aggaacaaga ggcgggcctt cgcagtgat tcaaagatgc cattgagatc 5640  
 gtagcggggt atgtttgtga tcacgccacc aatgaaaagc tgggtcaacaa ggcggtcgcc 5700  
 aggttcggac acctcgactg cttcatcgga aatgcgggga tatgggatta catgctgggc 5760  
 gtcgacgagc catgggagaa attatccggc agtttcgagg agatatttga tatcaacgtc 5820  
 aaaagctatt tcagcggcat cagcgcggcc ttgccggaac tcaaaaaaac gaacggatcg 5880  
 25gtcgtagtga cggttccgt ttcttctat gcggccggcg gcggcggtc ttgtacatt 5940  
 gccagcaagc atgcagtgtt gggatgtgtc aaggcgttgg cctacgaatt ggccccgcac 6000  
 atccgggtca atggcgttgc gccaggtggt acggctactt ctttggttgg gccggcaagc 6060  
 gccggtttcg acaaaaccaa aatgaaagac atgcccgga tcgatgacat gatcaagggc 6120  
 ctgacccctt tggggttcgc agcaaggccc gaggacgtgg tggcaccgta tcttttctg 6180  
 30gcctcccggt aacaaggga gttcatcact ggcaccgtaa tcggcattga tggcggcatg 6240  
 gcgctcggtc gaaagtgaat tttcaatcaa atcagatttt tcaaccccat tcccaggaga 6300  
 caacccatga agacgaaatt gttcatcaac aacacctgga gcgcttcgag tgacaaaaag 6360  
 tcattcgatc gcaagcacc tgtcagtggc gaggtcgtga cccaatgcgc gaacgccacg 6420  
 gtggacgatg cggccaatgc ggtcgcagcc gctcaagagg cgttcaagtc ctggaaggcc 6480  
 35gtcggaccct cggagcggcg gcgccttctt ttgaaggtgg cagacgtcat ggagagcaaa 6540  
 acgcccaggt tcacgaagt gatggccaag gaagtgggag cctccgcgct gtgggcgggg 6600  
 ttcaacgtgc acctgtcggc caatgtattc cgggaagccg cctcactggc cacccaaatt 6660  
 caaggcgaaa ccattccgac ggacaagcct gacaccctgt caatgacgct gcgtcacgct 6720  
 gtcggcccca tcttgagcat cgttccctgg aacggcaccg ccgtgctcgc ggcgcggggc 6780  
 40atcgcttacc cgctggtctg cggcaatacc gttgtgttca aaggctccga gttcagcccc 6840

	ggtacgcacg	ctttgatcac	caagtgcctg	caggaggcgc	acctgcctgc	tggcgtgctc	6900
	aactatctga	actccctccc	ggaccggctg	cccgatattg	cagatgcgct	gatttcgtct	6960
	aaagagattc	gtcgcacaa	cttcacaggc	tccactcgcg	tggggcgcat	catcgcccag	7020
	aaatcgcccc	agcatctcaa	gcgtgtcttg	ctggagttgg	gtggaagtc	cccgtgatc	7080
5	gtttctggacg	acgctgacat	cgacggcgca	gtcaaggcgc	cggtgttcgg	cagcttcctg	7140
	ttccaaggcc	agatctgcat	gtccacegaa	cgcctggtgg	tcgacgaaaa	gatcgcgga	7200
	gaatttgtcg	cgaagtctgt	cgagaaaacc	aagcagttga	gcgcaggcga	tccatgcgtc	7260
	acaggggact	gcatcatcgg	ccgatggtg	tcgcccgaact	cgggtgaccg	aatcaatggt	7320
	ctgttcaaag	atgccatcag	caagggcgcc	aaggtcgtgt	gcggcggcat	cgccgagggg	7380
10	gcggtcatgc	ccgccacgat	cttggaaccac	gtgacagcgc	acatgcagat	ctacgatgag	7440
	gaaaccttcg	gtcccatcac	tgtggttatc	cgggtgcaaga	gcgaagcggg	cgccatccgc	7500
	attgccaatg	acagcgccta	cggcctgtca	tcgggcgtgt	ttggccggga	cgtgaaccgg	7560
	gctctgcgcg	tgggcatggc	gatcgaatac	ggctcgggtc	atatcaacgg	ctccaccgta	7620
	cagaacgagg	ctcaggcgcc	ttatggcgga	acaaaggcca	ccggttatgg	ccgcttcgac	7680
15	ggacgcgcgg	tgatcgacga	gttcacggaa	ctcaagtggc	tgaccattga	accattcgag	7740
	cagcagtatc	ccttctaagc	tgaagcaaca	aaggagttaa	acctgaaca	agccagcaac	7800
	tgtcattgaa	ttgggggtaca	tgggcatttc	ggtcaaggat	ccgcgacgct	ggaaatcctt	7860
	tgccgcgaac	atgctgggac	ttcaagtcct	cgatgagggg	gacaaggatc	gcttctatct	7920
	gcgaatggac	aattggcacc	atcggatcgt	ggttcatcac	aacggtcaag	atgaccttga	7980
20	atacctgggc	tggcgtgtcg	ccggtcaacc	ggaattcgag	gcattgggtc	aaaagctcgt	8040
	ggacgcaggc	tacaaagtcc	gcgtgtgcga	caaagccgaa	gcacaagaac	ggatggtgct	8100
	gggcctgatg	aagacagaag	atccgggggg	caaccgcacc	gagattttct	ggggaccccg	8160
	gattgacctg	aacaaccctt	tccatcccgg	tcgtcccttg	cacgggaaat	ttctaaccgg	8220
	tgatcagggc	ctgggccact	gcacgtgcg	tcagaacgat	gttgaagcgg	cacgtaagtt	8280
25	ctatagcttg	ctgggatttc	gtggagatgt	cgagtaccgc	cttcctttgc	ccaacggcat	8340
	gacggctgag	ttgacgttca	tgcattgcaa	tgctcgcgat	catttcoatc	ctttcggtgc	8400
	aatgcctcgc	gccaaagcgc	tcaatcatct	gatgattgaa	tacactcata	tcgaagattt	8460
	gggttgacaa	caccagcttt	tcacgaagga	aaagattgac	attgccttgc	aattgggcat	8520
	ccattccaac	gataaggcgc	tgacgttcta	cggggcaaca	ccttcgggct	ggctgataga	8580
30	acctggggtg	cgaggcgccc	ccgccattgc	tgaatcggaa	tattacgtcg	gcgacatttt	8640
	cggccacacc	atcgaggcca	ccggttatgg	attggacgtc	aaactgagct	agccatgtaa	8700
	cagatgcgaa	atcgatcgca	tctgttttcc	ccattcagtt	ccatataaaa	aaaaggagac	8760
	aatgatgatc	aaagaagcca	tttccctcgc	aggacttggg	atgctgatgc	tcagcaccgc	8820
	atatgccgag	gattcccggt	ggtcatatcg	catcggcgc	accaacgtag	ccttcgatgc	8880
35	gagcgccaaa	gtttcgattg	acggaacaag	ggtgccgggt	ggaagcgtg	acgccagtga	8940
	caacaacgct	ttgacatttg	acttcggtta	catcattaac	gataactgga	acgcgcgatt	9000
	aattgttggc	attccaccca	ccacaaaagt	gacgggcgca	ggcacgctgc	ctccgatttt	9060
	gctgggcgct	gtccaatatg	ctcctgcagt	tttgtcggcg	acctacaacc	tgccacagat	9120
	gggatttggt	cgcccgtagc	tgggggcggg	gatcaactac	actcgaattc	tgaaaagcaa	9180
40	agatgccaat	ctgacctcat	tcgatgcaga	tcatgcgtgg	gcacttgtgc	tccacatcgg	9240

tcgagaagcg	aacatcaacc	gcgactgggt	cgtcagcttt	gatatccgaa	aactttatct	9300
aaaaacagat	gcatcggggt	ttcttggggc	tcaagttgct	acagcccgcg	taacgttgaa	9360
tccgctgctg	acgtcgattg	cgattggccg	gagattctga	tcggtccaca	ttgattccaa	9420
aattctgttt	gcatcaccat	tttcaaagga	aatttgaaatg	acaagaaaga	cgagcaaagc	9480
5ggtgcgcctg	accgccgcgg	atattcaagg	cgcattgggtc	atcatgccga	ccccgtccac	9540
gccggatgcc	tcggactggc	gcagcacgca	cacggtcgat	ctcgacgaga	cggccccgat	9600
tgtcgaaggag	ttgattgcgg	ccggcgtaaa	tggcattctg	agccacggca	cctttggcga	9660
atgcgcgacg	ctgacgtggg	aggagaagcg	ggattttgtt	tcaacggtcg	tggaaaccgc	9720
gcgcggtcga	gtgccctact	tctgcggcac	aacggccttg	aatacccggtg	aagtcatacg	9780
10ccagaccgcg	gaattgatcg	acattggcgc	ccaaggaaca	atgctcggcg	tgcgatgtg	9840
ggtgaagatg	gatctgccta	ctgccgtgca	attttatcgc	gatgtggcgg	aagcagtgcc	9900
agatgcagcc	atcgctgtct	acgccaaccc	ggaggctttc	aaatttgatt	ttcctcgccc	9960
gttttggggc	gaaatgtcca	aaatcccgc	ggtgtgcaca	gccaaagtact	tgggcacatcg	10020
gatgctggac	ttggatctga	aattggcccc	aaatattcgc	ttccttcgcg	atgaggatga	10080
15ctactacgct	gcggcccggg	tcaatcccga	gcgcattgact	gctttctggt	ctagcgggttc	10140
catgtgcggc	ccagcgaccg	cccttgctg	gcgcgatgag	gtggtaaagg	ccaaaaatac	10200
aggtgattgg	gccaaaggcca	aggctatttc	agatgacatg	cgcgacggcg	atgccacact	10260
gtttccacgc	ggcgatttct	cggaattctc	aaaatacaac	attggcctcg	aaaaagcacg	10320
aatggacgag	gccggctggc	tcaaggcggg	gccgtgcggg	ccaccctata	cgtctggttc	10380
20cgacgaatac	cttgacagtg	cccgaaaatc	aggcaaggcc	tggggccgcac	tgcataccaa	10440
gtatgccaa	gaattgagga	aaacaaaaac	ggcaaccaac	tcgaaaaaga	agtaagtcca	10500
ggccctgagt	cagacatctc	cgatcagcac	aacctgctga	tctggagggt	ttctggatta	10560
gtgcagtcgg	cgagttaaaa	atatgccagt	acagagtggg	agcgccacaa	gcgcaggggtg	10620
caccgatcgc	cctgaagcaa	tctcacacta	ttgccatctt	ctctttgttg	ggagtgcattg	10680
25atcgtgatgg	tcgattttta	tttcgatttt	ttgagcccat	tttcgtatct	ggccaaccac	10740
cgtttgtcgg	tgtctgcggg	gcgttatgga	ttctccatcc	agtatcacgc	cattgatattg	10800
gcgcgagcaa	aaacggccat	tggcaacatc	gggccatcca	atcgggacct	caaggccaag	10860
cttgactact	taaagtgga	tttgacgcga	tgggcccgatc	tctataggat	tccgttggtt	10920
ttccccctta	acttcaacag	ccgccgggtg	aatgccggac	tgtattaccc	ggcagccagg	10980
30gagcgagccg	ctgaatatgt	tcgccttggt	ttcgatttcg	cttgggggaa	aggggtgggca	11040
ctggatgctg	atagcttgct	ggctgaggta	tgcgacaagc	taaactggga	tctcggtgaa	11100
tttgaagatt	ttttgaacag	cgaaaatgcc	gccaaaggcat	acgacgaaga	gacgcaggcg	11160
gccattgacc	gaaagggttt	cgggggtccc	accgtgtttt	gggatgatca	aatgtggtgg	11220
ggaaatgacc	gccttttcat	gcttgagagc	aggttgcaaa	aggaacgcga	accataaatt	11280
35ctcagtgcatt	cgatttttct	gatttttcaa	ccacgccttg	attttaaggga	gtctcatgaa	11340
gctttattac	agccccggcg	catgttcac	gtgcctcac	atcatccttc	gtgaagggtgg	11400
atttgacttt	cagctagaaa	aggttgatct	cggcaccga	gtgactgaga	ctggtgttga	11460
ttacaagacc	gtcaatccc	tgggcagtg	ccccgcttg	caaatggatg	atgggcagggt	11520
gctcaccgaa	gggcccggcca	tcgttcaata	cctcgctgac	cgtgtgccag	aaaaatgcct	11580
40ggcaccggca	gctggctcgc	tggagcgtta	ccgactgatg	gaatggctga	atttcattctc	11640

	caccgaattg	cataaaaagct	tcggaagcgct	gttcagccccg	gtgtttccac	aggatgccaa	11700
	gccagtcac	aaggcccaat	tggaaagccg	tcttgcccat	accgagcaga	tgctcgggta	11760
	caaggtttg	gccatgggga	acgaactttc	cgtggtggac	gcctacttgt	ttactgtgct	11820
	tggctggggg	gcttacgtga	atgtggacct	ttcgcccttg	cccggcctgc	aaggctacct	11880
5	caaccgcgtg	gctgaacgtc	cagcagtcct	ggccacattg	tcggctgaag	gtttgatctg	11940
	atTTTTTTtag	agaaataaat	aaagcatggc	cagcctccat	gagcggccag	ccatgccgcg	12000
	gggcagccgt	gggcagcggc	ctcgatcaaa	tttcgatgca	cacagattac	gagccccctt	12060
	ttccgactgc	ggtcggattc	agaaattgcc	gatacttgct	ttggtctgat	cgaagcgtga	12120
	gggatatcgg	cttccccctat	ccacatgctt	gatgaagagg	agcgtatcac	catgagtcac	12180
10	gaacttggcc	gactgggaaga	cctgccgcag	gactaccgag	acgaactcaa	acaacttaac	12240
	ctggtgccc	tatggcccag	cctacgcgcg	gtgctgccgc	ccaatgtccc	gacccgccag	12300
	acgcagccga	cttactggtc	ctaccagacg	ctcaagccgc	tgctgtctaa	ggccgggtgag	12360
	ctgaccccc	ttgagaaggc	cgagcgcgcg	gtgctgggtg	tggccaaccc	cggccacggc	12420
	ctggagaaga	tgcaagccag	cgcgcgcata	tacctgggca	tgcagttgct	gctgcccgga	12480
15	gagtggggcg	ccagccaccg	ccacaccccc	aacgcgggtg	gcatgatcgt	ggagggcgag	12540
	ggcgccata	ccaccgtgga	tggcgagaag	tgccccatga	gccgggggtga	cctcatcttc	12600
	acgcccaccg	gcctgtggca	cgagcacggc	cacgacggca	acgagcccg	ggtgtggctc	12660
	gatgtgctg	acctgccgct	ggtgtactac	atggaggcca	gctaccacat	cgacggcgag	12720
	cgcacgcagg	tcgaccccgc	cgggggcgac	tgcgccctga	cccgtgcggg	cgtggtgcca	12780
20	acccccgtgt	tccagcgcag	cgacaagcgc	tatcctctct	tgcgctacce	ttggggccgac	12840
	acccgtgcgc	ccctgctgtc	gctggcggcc	gaccagccctg	agcaggaatg	cgtgcaggtc	12900
	acctacgtca	accccgaaac	gggtgacgac	gccgagaaca	tcttgggctt	ctacgccctg	12960
	atgctcaagc	ccggccagac	cctgcgcctg	cccgtgctgt	cgcgcccgct	ggtgttccac	13020
	cagatcgaag	gccgcagcga	ggcgcgcac	gccaggtcca	ccttcgcctt	gaggggaagc	13080
25	gatacctgct	gcgcccccg	ctacaccgag	gtgacgctga	aaaacctctc	agccgaccag	13140
	cgtccttca	tcttcatggc	cgacgaatcg	cccctgcacc	gcaagctggg	cgtctttgaa	13200
	aaccgcggct	gagccgcgcg	gaagcaccac	acaacaggaa	agcaagtggg	caactacctc	13260
	tggaaaccgc	ctcccgcca	gtctctgccc	gttcgtggca	agaccgagcg	cttccccatc	13320
	aaccgcatct	tctgcgtcgg	ccgcaactac	cacgcccacg	cgggtggaaat	gggcgctccg	13380
30	gtcgacaaaa	gcgtcgaaca	agcgttctac	ttacacaaat	cgcacaaaac	cctggtggaa	13440
	agtgggtcga	ccgtggccta	cccgcgcgcg	accagcaact	accactacga	gatggagctg	13500
	gtgctggcga	tcggcaagcc	cggcttccgc	gtcagcgaag	accaggcgca	cgagttgatt	13560
	tacggtctag	ccgccggcct	ggacatgacc	cggcgcgacc	tgcaactggg	ggcacgcgac	13620
	aagggccgcc	cctggggacac	cggcaaggac	atcgaggaag	gctcgggtctg	ctccgagatc	13680
35	gtgccgatgc	aaggcgtggg	gggtggagcag	ggcgcgatcg	ccctggaagt	caacggccag	13740
	accaaacagt	cgtccaacgt	ggacaagctg	atctggaacg	tccgcgagat	cattgccgac	13800
	ctctccacct	actaccactt	gcaacccggc	gacctgattt	acaccggcac	gcccgaaggc	13860
	gtgggcgcgc	tgggtggcggg	tgacaagatc	atcggcctgt	tggaaggcat	tgccgagatc	13920
	agcctgacgg	tcggcccggc	cgagtgagcc	tgacgcgat	gaagctgtac	aactcttggc	13980
40	gcagcggcac	gtcacaccgg	ctgcgcctcg	cgtctaacct	caagggcggtg	ccctacgaqt	14040

acctggccgt gcacctgggc aaggaagagc acctaaagga cgccttcaag gcgctgaacc 14100  
 cacagcagtt ggtgcccgcg ctggacacgg gcgcgcaagt gctgatccag tcgccggcca 14160  
 tcatcgaatg gctggaggaa cagtacccca cgcccgcgct gctgccggcc gacgccgacg 14220  
 gccgccaacg ggtgcgcgcg ctggccgcca tcgtgggctg cgacatccac cccatcaaca 14280  
 5accgcccgtat tctggagtac ctgcgcaaga cgttcggcgc cgacgaggcc gccatcaacg 14340  
 cctggtgttg cacctggatc agcgccggtt tcgatgccta cgaagcactg ctggcggtcg 14400  
 acccgaagcg cgcccgctac agctttggcg acacaccac gctggccgac tgctacctgg 14460  
 tgccgcaggt ggaaagcgcc cgccgcttcc aggtggacct gacgccctac cccctgatcc 14520  
 gcgcagtgga cgcgccctgc ggcgagttgg acgcatttcg gcgcgcgcg ccagctgctc 14580  
 10aacctgattc ggcttgacga aaaatacgcc gccttcgccc taataaacag ttcgtagccc 14640  
 tgcccagtga caagtatcaa tgagcaatag tgaaaatgta aggggaaaga ttatgaaaca 14700  
 tattctgacc cgacgagcag ccctgagcac tttgggctcc cttggaatcg ggggccttcc 14760  
 gggcatgagc ttggcgcaat ccggccctgg cgtggccacg atcgtgattg gactcgccgc 14820  
 aggcggagcg accgatatgg ctgcccgccg gttgtctgag ggaatgcgcg gtgcctatgc 14880  
 15ttccagtgtg ctgctcgaca accgcacggg ggccgggtgtg cgaattgcga tccagcatgt 14940  
 taaggccgcg gcaccaaacg gcgcgacact tctgctgacg cctgcattcca tgatgacgct 15000  
 ataccctcac acgtacaaa atct 15024

20

<210> 12  
 <211> 4912  
 <212> DNA  
 <213> Artificial Sequence

25

<220>  
 <223> A modified DNA sequence

&lt;400&gt; 12

30 gagctcgttg cgcaagcgt tccatcagtg ggggcatgaa gtacatctcg ccgatgtcgg 60  
 tcattgccaa gttgaagggt gcgctgctgg caaatgggtc gaaagagtca cgggtcgtca 120  
 gtgcgctctg cagcgtgttg agcgcataga ccacgggctc cgcaagatgc agtgcatacg 180  
 gtgtcggttc catgcctttt gaggtgcgca agaacaaatc gtccttttagc gccgcacgca 240  
 gccgtttaag tgaattgctg acggcaggct gcgtcagccc cagtttttcg ccggcgctcg 300  
 35 atacgctccg gtcgagcagt agctggttga agaccaccag cagattcaag tcgatgtcgc 360  
 gcagatccat gatgcctcac cattattcat gctggtgatt ttaactatca gacttgatct 420  
 atagcgctat accgatcgac gcgccagtat cgcagccatt cggagacaac tgaaaaaaga 480  
 gcttgcatgg aactggtagt agaaccctc aatttgcatc tgaacgcgga gaccggcagc 540  
 accctgcttg acgtgctcag gtccaacgag gtccccattt cttatagctg catgtcgggc 600  
 40 cgctgcggca cttgccgttg ccgtgtgatt gccggccatc ttcgcgataa cggtccgag 660

	acagggcgcc	cgcaggcagg	aaagggggcc	tatgtcctgg	cctgtcaggc	ggttctgacc	720
	gaagactgca	cgatcgagat	tcttgaatct	gacgagatcg	tggttcaccc	ggcgcgcac	780
	gtcaagggga	cggtcacagc	gatagacgaa	gccacccatg	acatccggcg	cctgcgcac	840
	aaactggcca	aaccgcctta	gttcagccct	ggccagtagc	caacgggtgca	gttcacgccc	900
5	gaatgcgtcc	gccccatttc	gatggccggg	ctgcctagcg	atgcgggaaat	ggagtttcag	960
	attcgcgcgg	ttccggggcg	gcatgtcagc	aactacgttt	tcaatgaact	gtccgtaggc	1020
	gcttcgggtgc	ggatcagcgg	ccccctcgga	acggcctatc	tgcggcgcac	gcacaccggc	1080
	cccattgcttt	gtgtgggggg	tggaaçaggt	ctggcgcgcc	tcttttcgat	cgttcgaggc	1140
	gcactggaaa	gcgggatgag	caaccccatc	catctgtact	tcgggtgtgcg	gagcgagcag	1200
10	gacatctatg	acgaggaacg	ccttcacgca	ttggctgcaa	ggtttccgaa	tctcaagggtg	1260
	aatgtcgttg	ttgcaacagg	ccttgccggc	cctggtcac	gatccggcct	ggtcaccgat	1320
	ctgatcggcc	gtgacttgcc	caatttgggc	ggatggcgcg	cctacctgtg	tggcgctccg	1380
	gccatgggtcg	aggccctgaa	cctgctcgtt	gctcgcctag	gcatagtacc	cgggcacatc	1440
	catgccgatg	cgttctatcc	cagcggcgtc	tgagcgaaag	caccatgcga	acccaattca	1500
15	acccaaggat	accaagccat	gagtgaaccc	caacgattaa	aaccctgtgt	tccccaagat	1560
	ccgaaatggc	cgggcgaagg	tagcagccgc	gttcccttct	gggcctacac	ccgcgaagac	1620
	ctgtacaagc	gcgaattgga	gcgcctgttc	tatgcaaacc	actggtgcta	tgtaggcctg	1680
	gaagccgaga	ttccgaatcc	aggcgacttc	aagcgaaagg	tgatcgggtg	gcgctcggtc	1740
	atcatgggtgc	gtgatccgga	tggcggcatc	aacgtggtgg	agaacgtctg	cgcctactgt	1800
20	ggcatgcgct	tttgcccgga	gcgccacggc	aacgccaagg	acttcttctg	cccctaccac	1860
	cagtggaaact	acagcctcaa	gggtgacctg	cagggcgctg	ccttccgcgc	aggcgtaag	1920
	caggacggca	aggtaacagg	gggcattgcc	aaggacttca	aactcgaaga	acacggcctg	1980
	accaagctca	agggtggcgc	ccgaggcggt	gcagtgtttg	cctctttcga	ccacgatgtc	2040
	gagcctttcg	aggactgtgc	gtgtacgaca	gcgagatgat	ccccaaactc	ctcatctacc	2100
25	ccatttgagg	ttgcagaaca	tgtccgagaa	ctggattgat	gccatcgcac	gggacgctgt	2160
	gcctgagggg	gatgtggtcg	gagtcacgt	ggcaggcaaa	gacattgcct	tctatgaggt	2220
	ggaagggtgag	gtcttcgcc	ccgacaactt	gtgtaccac	ggggctgcgc	gcttgagcga	2280
	cggcttttct	gaaggccggg	aaattgaatg	tcctttgcat	caaggccgat	tcgatgtttg	2340
	cacgggtaaa	gccttggtga	cacccttgac	acaggacatc	aaaacctacc	ccgtaaaaat	2400
30	cgaaaacatg	cgcgtgatgc	tcaagctgga	ctaaaactct	ttgcaggagg	aaagccaaat	2460
	ccggaaatca	ccccacccaa	cccaatcact	acccgttttc	aaacaagagg	agataagcaa	2520
	ttatgagtta	ccaaaactta	gtgagtgaag	cagggctgac	gcaaaagctc	ctgattcatg	2580
	gcgacaaaga	acttttccag	cacgaattga	agaccatctt	cgcgcggaac	tggctttttc	2640
	tgacccatga	cagtctgatt	ccctcccccg	gcgactatgt	cacagccaaa	atgggcgtcg	2700
35	atgaagtcac	cgtctcccg	cagaacgatg	gctcgggtgc	agcctttttg	aatgtttgcc	2760
	gtcaccgggg	caagacacta	gttcacactg	aagccggaaa	tgcgaaaggc	tttgtgtgcg	2820
	gtaccacgg	ctggggctac	ggttccaacg	gcgaactgca	aagcgttccc	tttgaaaaag	2880
	agttgtacgg	agatgcgac	aaaaagaaat	gcctgggctt	gaaagaagtc	ccccgcacg	2940
	aaagctttca	tggctttatc	tatggctgtt	ttgatgcaga	agctcccccg	ctcatcgatt	3000
40	atctgggtga	tgcagcctgg	tacctggaac	ccaccttcaa	gcactctggt	ggcctggaac	3060

ttgtaggccc ccccgcaaaa gtggtggtta aggcactg gaagcctttt gcggaaaact 3120  
 ttgtagggtga catctaccac gttggttga cgcacgcagc ggctttgcgc gcagggcagt 3180  
 cggatatttag ttctcttgcg ggcaacgcta agcttccacc cgaaggcgcg ggcttgcaaa 3240  
 tgaccagcaa gtatggcagt ggaatgggct taacgtggga ctactactcc ggtaacttca 3300  
 5 gcgctgatat ggttcccgat ctgatggcat tcggcgccgc aaaacaggaa aaactcgcca 3360  
 aggaaatcgg cgatgtccgg gcacggattt accgcagcat tctgaacggc acggttttcc 3420  
 cgaacaacag ctttttgacc ggctccgcta cttcaaggt ctggaaccgc atcgatgaaa 3480  
 acacgaccga ggtttgagc tatgccttcg tagaaaaaga catgcctgag gacttaaagc 3540  
 gtcgcttggc tgacgcggt cagcgagtg tcggaccagc aggattctgg gaaagcgagc 3600  
 10 acaacgaaaa catggagacg ttgtcgcaaa atgccaagaa ataccaatcc agcaacagt 3660  
 atcagattgc cagtttgggt ttccgcaagg acgtctacgc cgacgaatgc tatccggcg 3720  
 tcggttgcaa atcggcaatc ggcgaaacca gctatcgcg attctaccgt gcctaccagg 3780  
 ctacatcag cagctccaat tgggcccagtc tcgaaaatgc ctcccgaaat tggcacaccg 3840  
 aactcaccaa gacaactgat cgctaacca ggagccaacc atgatgatca ataccagga 3900  
 15 agacaagctg gtctccgcgc acgacgccga agaatttcac cgtttcttca tcgtacaaga 3960  
 tgatgcacta ctgcaagaag tcaacacgct cctgaccgc gaagcgacc tgctggacat 4020  
 tcaggcctac aaagcctggc ttgaacactg cgttgcccc gagatcaaat accaagtgat 4080  
 ctccgcgagaa ctctcctcca ctcccgagc tcgataccaa ctgaatgatg cggatgaatct 4140  
 ctacaacgag aactatcaac agctgaaagt tcgagtgaa caccagatgg atcctcagaa 4200  
 20 ctgggccaac aaccggaaga tccgcttcac ccgcttcgtc accaatgtca cggcgcccaa 4260  
 ggacaagagc gcaccggaat tgctgcatgt ggggtccaac ctcatctctc atcgcgccag 4320  
 acgagaaaac caagttgagc tcttctatgc aacgcgtgaa gacaaatgga aacgcacga 4380  
 aggtggtggt atcaaattgg tcgaacgctt tgtggactac ccggagcgca tccccagac 4440  
 ccacaacctg ttggtcttcc tgtgaacctt ggggatgcct gcctggatgg cgggcattcg 4500  
 25 tgattatttt taacagaaat ttattgcat gaacacacag caagttgttg ccatcactgg 4560  
 cgcggctcg ggcattggt tcgagttggt tcgctctttt aaggcagccg gttatcgct 4620  
 atctgcactc gtccgcaacg aggagcaaga ggcgggtctt cgcagtgaat tcaaagacga 4680  
 cattgagatc gtggcgggcg atgtccgtga tcacgccacc aatgagaagc tggtaaaca 4740  
 gacggttgcc aagttcgggc gcctggattg cttcatcgga aatgccggga tatgggatta 4800  
 30 catgctgagc atcgatgatg cttgggagaa attctcgggc agtttcgagc agatatttga 4860  
 catcaacgctc aaaagctatt tcagcggcat cagcgcgcc ttgccggagc tc 4912

<210> 13

<211> 6779

35 <212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

40

&lt;400&gt; 13

atgcatacgt agcgggtggcg aaagaggcgc cgctggaaga atccggctctg tagctctggc	60
agaagaccca aggcgagatc aaccgcaccg gactccatat cctccttcag attgccagca	120
ttcggggcgca gcgtgctgat ctggatgtga ggagctcggt gcgcaagcgc ttccatcagt	180
5gggggcatga agtacatctc gccgatgtcg gtcattgccca agttgaagggt gcgcgtgctg	240
gcaaatgggt cgaaagagtc acgggtcgtc agtgccgtct gcagcgtgtt gagcgcatag	300
atcacgggct ccgcaagatg cagtgcatac ggtgtcggct ccatgccttt tgagggtgcgc	360
aagaacaaat cgctcgttcag cgctgtacgc agccgtttta gtgaattact gacggcaggc	420
tgcgtcagcc ccagtttttc gccggccgtc gatacgtcc ggtcgagcag tagctgggtg	480
10aagaccacca gcagattcaa gtcgatgtcg cgcagatcca tgacctacc ctcaccatta	540
ttcatgctgg tgattttaac tatcagactt gatctatagc gctataccga tcgacgcgt	600
agaatcgag ccattcggag acaactgaaa aaagagcttg catggaactg gtagtagaac	660
ccctcaattt gcactgaac gcggagaccg gcagcaccct gcttgacgtg ctcagggtcca	720
acgaggtccc catttcttat agctgcatgt cgggcccgtg cggcacttgc cgttgccgtg	780
15tgattgccgg ccactctcgc gataacggcc ccgagacagg gcgcccgcag gcaggaaagg	840
gggcctatgt cctggcctgt caggcgggtc tgaccgaaga ctgcacgac gagattcctg	900
aatctgacga gatcgtggtt caccgggcgc gcactcgtcaa ggggacggtc acagcgatag	960
acgaagccac ccattgacatc cgggcgcctgc gcacaaaact ggccaaaccg cttgagttca	1020
gccttgccca gtacgcaacg gtgcagttca cggccgaatg cgtccgcccc tattcgatgg	1080
20ccgggctgcc tagcgatgcg gaaatggagt ttcagattcg cgcggttcg ggccggcatg	1140
tcagcaacta cgttttcaat gaactgtccg taggcgttc ggtgcggatc agcggcccc	1200
tcggaacggc ctatctggcg cgcacgcaca ccggccccat gctttgtgtg gggggtgaa	1260
caggctctggc gcccgctcct tcgatcgttc gaggcgcact ggaaagcggg atgagctacc	1320
ccatccatct gtacttcggg gtgcggagcg agcaggacat ctatgacgag gaacgccttc	1380
25acgcattggc tgcaagggtt ccgaatctca aggtgaatgt cgttggtgca acaggccctg	1440
ccggccctgg tcatcgatcc ggcctggta ccgatctgat cggccgtgac ttgcccaatt	1500
tggcgggatg gcgactgcat cctgtgtggc gctccggcca tggtcgaggc cctgaacctg	1560
ctcgttgctc gcctaggcat agtaccggg cacatccatg ccgatgcgtt ctatccagc	1620
ggcgtctgag cgaaggcacc atgcgaaccc aattcaaccc aaggatacca agccatgagt	1680
30gaaccccaac gattaaaacc cgtgtttccc caagatccga aatggccggg cgaaggtagc	1740
agccgcgttc ccttctgggc ctacaccgc gaagacctgt acaagcgcga attggagcgc	1800
ctgttctatg caaaccactg gtgctatgta ggccctggaag ccgagattcc gaatccaggc	1860
gacttcaagc gaacggtgat cggtagcgcg tcggatcatca tgggtcgtga tccggatggc	1920
ggcatcaacg tggtagagaa cgtctgcgcc caccgtggca tgcgcttttg ccgcgagcgc	1980
35cacggcaacg ccaaggactt cttctgcccc taccaccagt ggaactacag cctcaagggt	2040
gacctgcagg gcgtgccctt ccgcccgtggc gtcaagcagg acggcaagggt caacggcggc	2100
atgcccaagg acttcaaac cgaagaacac ggccctgacca agctcaagggt ggccgccccg	2160
aggcgggtgca gtgtttgcct ctttcgacca cgatgtcgag ctttcgagg acttcctggg	2220
ccaaccatcc tgcactactt cgatcgcgtc ttcaatggcc gcaagctcaa gatcctgggc	2280
40taccgccgcc agcgcacccc gggcaactgg aagctgatgc aggagaacat caaggacccc	2340



	taccacccgg	gcctgctgca	cacctggttc	tcgaccttcg	ggctctggcg	cgcgcagaaac	2400
	aagtcggaac	tgaagatgga	cgccaagttc	cgccacgcgc	caatgatctc	cacgcgcggt	2460
	cagggcgga	agaacgagga	ggtcgtgtcc	ggcgtggaca	gcttcaagga	acagatgaag	2520
	gtgaacgacc	cgcgcctgct	cgacatcgtg	cccgagccct	ggtggggcgg	tccgactgcg	2580
5	gtgatgacca	cgatcttccc	cagcgtgata	atccagcagc	aggtaaacag	cgtatcgacc	2640
	cgccacatcc	agcccaacgg	tcacggctcc	ttcgatttcg	tctggaccca	cttcggcttc	2700
	gaggacgaca	acgaggagtg	gaccacgcgc	cgcctgatcc	aggccaacct	gttcggggccg	2760
	gcgggcttcg	tgtcggccga	tgacggcgag	gtgatcgagt	ggtcgcagga	aggctttgag	2820
	caaaaaccga	cgcaccgcac	cgtgatcgag	atgggcggtc	acgaaatcgg	cgacacggac	2880
10	cacatggtca	ccgagacgct	gatccgcggc	atgtacgact	actggcgcaa	ggtgatgggg	2940
	gaataaacat	ggtagacttc	aaaacctatt	tcgaactgct	gaacctgtac	agcgactacg	3000
	ccatggtgtg	cgactccgcc	aattgggaga	agtggcctga	tttcttcata	gagaccggca	3060
	cctaccgctc	gcaaccgcgc	gaaaacttcg	agcaggactt	gccgctgtgt	ctgctggcgc	3120
	tggagagcaa	ggccatgatt	cgtgaccgag	tgtacggtgt	caaggaaacc	atgtaccacg	3180
15	atccctacta	ccagcgccac	atcgtaggca	cgcgcgcgct	gctgtcagtg	gagcgtgatg	3240
	cggacggcga	gcgcatcacc	gccgaagcca	gctatgccgt	gattcgcacc	aagtacgacg	3300
	gcgattccac	gattttcaac	gccggctatt	accgagacgt	gatcgtgcgc	acgcccgagg	3360
	gcctcaagct	gaagtcgcgc	ctgtgcgtgt	acgacagcga	aatgattccc	aactccatca	3420
	tctaccctat	ctgagaagga	atccaatgag	cgagaactgg	atcgacgcgc	ccgcccgcga	3480
20	cgaggtgcca	cgagggcgac	gtgatcggca	tcaatatcgt	cggcaaggag	attgcctcta	3540
	cgaggtgggc	ggcgagatct	acgccaccga	caacacctgc	actcacggcg	ccgcccgcac	3600
	gagcgatggc	tttctcgaag	gccgggaaat	tgaatgtcct	ttgcatcaag	gccgattcga	3660
	tgtttgcacg	ggtaaagcct	tgtgcacacc	cctgacacag	gacatcaaaa	cctaccccg	3720
	aaaaatcgaa	aacatgcgcg	tgatgctcaa	gctggactaa	atgctcaagc	tggactaaaa	3780
25	ctcttttcag	gaggaaagcc	aaatccggaa	atcaccccac	ccaacccaat	cactaccgct	3840
	tttcaaacaa	gatgagacaa	gcaattatga	gttaccaaaa	cttagtgagt	gaagcagggc	3900
	tgacgcaaaa	gcacctgatt	tatggcgaca	aagaactttt	ccagcacgaa	ttgaagacca	3960
	tcttcgcgcg	gaactggctt	tttctgacct	atgacagtct	gattccctcc	cccggcgact	4020
	atgtcaaaagc	caaaatgggc	gtcgatgaag	tcacgtctct	ccgccagaa	gatggctcgg	4080
30	tcgcgagcctt	tttgaatgtt	tgccgtcacc	ggggcaagac	aatagttgac	gctgaagccg	4140
	gaaatgcgaa	aggctttgtg	tgcggttacc	acggctgggg	ctatggctcc	aacggcgaa	4200
	tgcaaagcgt	tccctttgaa	aaagagttgt	acggagatgc	gatcaaaaag	aaatgcctgg	4260
	gcttgaaaga	agtcccccgc	atcgaaagct	ttcatggctt	tatctatggc	tgttttgatg	4320
	cagaagctcc	cccgtctatc	gattatcttg	gtgatgtagc	ctggtacctg	gaacccacct	4380
35	tcaagcactc	tgggtggcctg	gaacttgtag	gccccccgcg	caaagtgggtg	gttaagggca	4440
	actggaaggt	ttttgcggaa	aactttgtag	gtgacatcta	ccacattggt	tggacgcacg	4500
	catctatttt	gcgcgcaggg	caggcgatat	ttgctcctct	tgcgggcaac	gctatgcttc	4560
	cacccgaagg	cacgggcttg	caagcgacca	ccaagtatgg	cagtggaaat	ggcgatcgt	4620
	tggacgccta	ctccggtgtc	cagagcgctg	atctgggttc	cgaaatgatg	gcattcggcg	4680
40	gcgcaaaaaca	ggaaaagctc	gccaaagaaa	tcggcgatgt	ccgggcgcgg	atttaccgca	4740

gccaaagtga cggcacgggt ttcccgaaca actgcttttt gaccggcgcc ggtgttttca 4800  
 aggtcttttaa cccgatcgat gaaaacacga ccgaggttg gacgtatgcc atcgtagaaa 4860  
 aagacatgcc tgaggactta aagcgtcgct tggctgacgc ggctcagcgc tctgtcggac 4920  
 cagcaggata ctgggaaagc gacgacaacg acaacatggt gttgtcgcaa aatgccaaaga 4980  
 5aataccaatc cagcaacagt gatctgattg ccgatttggg ttccggcaag gacgtctacg 5040  
 gcgacgaatg ctatccgggc gtcgttagca aatcggcatt cagcgaacc aaccatcgcg 5100  
 gattctaccg tgccctaccg gtcacatca gcagctccaa ttgggcccag ttcgaaaaca 5160  
 cctcccgaaa ttggcacacc gaactcacca agacgactga tcgctaatac aggagccaat 5220  
 catgatgate aatacccagg aagacaagct ggtctccgcg cagcagccg aagaatttca 5280  
 10ccgttttctt gtcgggcaag acagcgatct gcagcaagaa gtcaccacac tctgacccg 5340  
 cgaagccgac ctgctggaca ttcaggccta caaagcctgg cttgaacact gcgttgcccc 5400  
 cgagatcaaa taccaagtga tctcgcgaga acttcgctcc acttcgagc gtcgatacca 5460  
 actgaatgat gcggtgaata tctacaacga gaactatcaa cagctgaaag ttcgagttga 5520  
 acaccagatg gatcctcaga actggtacaa cagcccgaa atccgcttca cccgcttcgt 5580  
 15caccaatgtc acggcggcca aggacaagag cgcaccgaa atgctgcatg tcggttccaa 5640  
 cctcattctc catcgcgcca gacgaggaaa ccaagttgac gtcttctatg caacgcgaga 5700  
 agacaaatgg aaacgcatcg aaggtggtgg catcaaattg gtcgaacgct ttgtggacta 5760  
 cccggagcgc agtcccaaaa cccacaacct gatgatcttc ctgtgagccc tggggatgcc 5820  
 tgcttgatg gcgggcattc gtgattattt ttaacagggc ggattcaaaa gtgaagtga 5880  
 20acaccctgga tttcagtga tgagagtga gtgctgccc ctattcaca gcagttcgcg 5940  
 gtagaccgcc agcggtgatc gtactcccag ccctttcctt ggccggttgt tgatttcac 6000  
 ggcaatcgca tcgagctgct cctggctgta gatgctcaga tccgtccctt tgggcaggt 6060  
 ctggcgccacc aggcattca tgttctcgtt ggagcctctt tgccaagggc tgtgcgggtc 6120  
 gcagaagtac acggctatgc cggtttgctc actgagcttc ttgtgcatgg acatctcccg 6180  
 25gccttggtcg tacgtcatgc tcaaccgat cggctgcgca atgccagca gcttgtccgt 6240  
 gaaggcctgc aagacgttg cagcactggc cggcttgaac tcaggcagct tgaccagcat 6300  
 caccagccgg ctggtgcggt ccaccaaggt cccgactgcg ctggcattgc cttctccctt 6360  
 gatgagatca cttcccagt gcccggaata ctggcggtct tcgatctcgg gcgggctac 6420  
 atggatgctc agcatgtcgg gaactctggc tctgcggtcc ttgcccttgc tgcgcggcag 6480  
 30ccgcttggtg tggcggtgac gcagggtggc gatcagctct ttgcgcagct ctcccacggg 6540  
 catggcatag atgcagttgt agatggttct gtgtgacacg cggatttcat ggcccagggg 6600  
 atacaaacgt gccagtgtca gggcaatctg ctgcggcgac cagcgcagcc gcagcagatg 6660  
 gatcaccagg caggacaaga tcgactcggg atgcagcttg atagctggtc tgccacagcg 6720  
 tctgcgctgc agggactggc actgggcatg tgcgctgcca tagccagatg aagatgcat 6779

35

&lt;210&gt; 14

&lt;211&gt; 449

40

&lt;212&gt; PRT

<213> Artificial Sequence

<220>

<223> A polypeptide encoded by SEQ ID NO:3

5

<400> 14

```

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
  1             5             10             15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
10             20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
  35             40             45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
  50             55             60
15 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
  65             70             75             80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
  85             90             95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
20             100            105            110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
  115            120            125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
  130            135            140
25 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
  145            150            155            160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
  165            170            175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
30             180             185             190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
  195            200            205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
  210            215            220
35 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
  225            230            235            240
Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
  245            250            255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
40             260             265             270

```

35

Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 50                      55                      60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65                      70                      75                      80  
 5 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
                     85                      90                      95  
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
                     100                      105                      110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 10                      115                      120                      125  
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
                     130                      135                      140  
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145                      150                      155                      160  
 15 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
                     165                      170                      175  
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
                     180                      185                      190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 20                      195                      200                      205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
                     210                      215                      220  
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
 225                      230                      235                      240  
 25 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
                     245                      250                      255  
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
                     260                      265                      270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 30                      275                      280                      285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
                     290                      295                      300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305                      310                      315                      320  
 35 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
                     325                      330                      335  
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val  
                     340                      345                      350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 40                      355                      360                      365

Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 5 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 10 435 440 445  
 Arg

<210> 16  
 15 <211> 449  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 20 <223> A polypeptide encoded by SEQ ID NO:5

<400> 16  
 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln  
 1 5 10 15  
 25 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys  
 20 25 30  
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 35 40 45  
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 30 50 55 60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65 70 75 80  
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
 85 90 95  
 35 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 115 120 125  
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 40 130 135 140

40

&lt;210&gt; 17

&lt;211&gt; 449 .

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

5

&lt;220&gt;

&lt;223&gt; A polypeptide encoded by SEQ ID NO:6

&lt;400&gt; 17

10 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Phe Gly Leu Ser Gln  
     1                    5                    10                    15  
   Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys  
                     20                    25                    30  
   Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 15          35                    40                    45  
   Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
           50                    55                    60  
   Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
   65                    70                    75                    80  
 20 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
                     85                    90                    95  
   Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
                     100                    105                    110  
   Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 25          115                    120                    125  
   Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
           130                    135                    140  
   His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
   145                    150                    155                    160  
 30 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
                     165                    170                    175  
   Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
                     180                    185                    190  
   Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 35          195                    200                    205  
   Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
           210                    215                    220  
   Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
   225                    230                    235                    240  
 40 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly



245 250 255  
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 260 265 270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 5 275 280 285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 290 295 300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305 310 315 320  
 10 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
 325 330 335  
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 15 355 360 365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 20 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 Tyr Arg Ala Tyr Arg Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 25 435 440 445  
 Arg

<210> 18

30 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

35 <223> A polypeptide encoded by SEQ ID NO:7

<400> 18

Met Asn Tyr Lys Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Thr Gln  
 1 5 10 15  
 40 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Arg

	20	25	30
	Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile		
	35	40	45
	Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val		
5	50	55	60
	Ile Val Ser Arg Gln Ser Asp Gly Ser Ile Arg Ala Phe Leu Asn Val		
	65	70	75
	Cys Arg His Arg Gly Lys Thr Leu Val Asn Ala Glu Ala Gly Asn Ala		
	85	90	95
10	Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly		
	100	105	110
	Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu		
	115	120	125
	Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe		
15	130	135	140
	His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met		
	145	150	155
	Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His		
	165	170	175
20	Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys		
	180	185	190
	Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His		
	195	200	205
	Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe		
25	210	215	220
	Ala Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu		
	225	230	235
	Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly		
	245	250	255
30	Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe		
	260	265	270
	Gly Gly Ser Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg		
	275	280	285
	Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn		
35	290	295	300
	Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp		
	305	310	315
	Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met		
	325	330	335
40	Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val		

145                      150                      155                      160  
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
                                  165                      170                      175  
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
 5                      180                      185                      190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
                                  195                      200                      205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
                                  210                      215                      220  
 10 Cys Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
                                  225                      230                      235                      240  
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
                                  245                      250                      255  
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 15                      260                      265                      270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
                                  275                      280                      285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
                                  290                      295                      300  
 20 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
                                  305                      310                      315                      320  
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
                                  325                      330                      335  
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val  
 25                      340                      345                      350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
                                  355                      360                      365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
                                  370                      375                      380  
 30 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
                                  385                      390                      395                      400  
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
                                  405                      410                      415  
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 35                      420                      425                      430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
                                  435                      440                      445  
 Arg

40

&lt;210&gt; 17

&lt;211&gt; 449

&lt;212&gt; PRT

	85	90	95
	Lys Gly Phe Val Cys Ser Tyr His	Gly Trp Gly Phe Gly Ser Asn Gly	
	100	105	110
	Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu		
5	115	120	125
	Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe		
	130	135	140
	His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Ser Leu Met		
	145	150	155
10	Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His		
	165	170	175
	Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys		
	180	185	190
	Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His		
15	195	200	205
	Val Gly Trp Thr His Ala Ser Ser Leu Arg Thr Gly Glu Ser Ile Phe		
	210	215	220
	Ser Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu		
	225	230	235
20	Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly		
	245	250	255
	Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe		
	260	265	270
	Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Pro		
25	275	280	285
	Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn		
	290	295	300
	Ser Val Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp		
	305	310	315
30	Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met		
	325	330	335
	Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val		
	340	345	350
	Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr		
35	355	360	365
	Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Ile		
	370	375	380
	Ser Asn Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Ala Val Tyr Pro		
	385	390	395
40	Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe		
		400	

405 410 415  
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu Asp Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 5 435 440 445  
 Arg

<210> 20  
 10 <211> 449  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 15 <223> A polypeptide encoded by SEQ ID NO:9

<400> 20  
 Met Asn Tyr Lys Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Thr Gln  
 1 5 10 15  
 20 Lys His Leu Ile His Gly Gly Glu Gly Leu Phe Gln His Glu Leu Arg  
 20 25 30  
 Ala Val Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 35 40 45  
 Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 25 50 55 60  
 Ile Val Ser Arg Gln Ser Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65 70 75 80  
 Cys Arg His Arg Gly Lys Thr Leu Val Asn Ala Glu Ala Gly Asn Ala  
 85 90 95  
 30 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu  
 115 120 125  
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 35 130 135 140  
 His Gly Phe Ile Tyr Ala Cys Ile Asp Gln Glu Ala Pro Ser Leu Met  
 145 150 155 160  
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His  
 165 170 175  
 40 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys

	180	185	190
	Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His		
	195	200	205
	Val Gly Trp Thr His Ala Ser Ser Leu Cys Thr Gly Glu Ser Ile Phe		
5	210	215	220
	Ser Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu		
	225	230	235
	Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly		
	245	250	255
10	Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe		
	260	265	270
	Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg		
	275	280	285
	Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn		
15	290	295	300
	Ser Val Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp		
	305	310	315
	Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met		
	325	330	335
20	Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val		
	340	345	350
	Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr		
	355	360	365
	Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Ile		
25	370	375	380
	Ser Asn Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Ala Val Tyr Pro		
	385	390	395
	Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe		
	405	410	415
30	Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe		
	420	425	430
	Glu Asp Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp		
	435	440	445
	Arg		
35			

&lt;210&gt; 21

&lt;211&gt; 449

&lt;212&gt; PRT

40 &lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; A polypeptide encoded by SEQ ID NO:10

&lt;400&gt; 21

```

5 Met Asn Tyr Lys Asn Lys Asn Leu Val Ser Glu Ser Gly Leu Thr Gln
  1             5             10             15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln Arg Glu Leu Glu
  20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
10             35             40             45
Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val
  50             55             60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65             70             75             80
15 Cys Arg His Arg Gly Lys Thr Leu Val His Ala Glu Ala Gly Asn Ala
  85             90             95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ala Asn Gly
 100            105            110
Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ala Leu
20            115            120            125
Asp Lys Lys Cys Met Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 130            135            140
His Gly Phe Ile Tyr Gly Cys Phe Asp Glu Glu Ala Pro Ser Leu Lys
145            150            155            160
25 Asp Tyr Met Gly Asp Ala Gly Trp Tyr Leu Glu Pro Met Phe Lys His
 165            170            175
Ser Gly Gly Leu Glu Leu Ile Gly Pro Pro Gly Lys Val Ile Ile Lys
 180            185            190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Thr Gly Asp Ala Tyr His
30            195            200            205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Gln Ser Val Phe
 210            215            220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225            230            235            240
35 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245            250            255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260            265            270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Glu Val Arg
40            275            280            285

```

Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 290 295 300  
 Ser Phe Leu Thr Cys Ser Gly Val Phe Lys Val Trp His Pro Ile Asp  
 305 310 315 320  
 5 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Met Val Glu Lys Asp Met  
 325 330 335  
 Pro Glu Asp Leu Lys Arg Arg Leu Val Asp Ala Val Gln Arg Thr Val  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 10 355 360 365  
 Val Ser Gln Asn Ala Lys Lys Tyr Gln Ser Arg Asp Gly Asp Leu Val  
 370 375 380  
 Ser Asn Leu Gly Phe Gly Gly Asp Val Tyr Gly Asp Glu Val Tyr Pro  
 385 390 395 400  
 15 Gly Ile Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 Tyr Arg Ala Tyr Gly Ala His Ile Ser Ser Ser Ser Trp Ala Glu Phe  
 420 425 430  
 Glu Asp Val Ser Lys Asn Trp His Thr Glu Leu Ala Lys Thr Thr Asp  
 20 435 440 445  
 Arg

<210> 22  
 25 <211> 447  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 30 <223> A polypeptide encoded by SEQ ID NO:11

<400> 22  
 Met Ile Tyr Glu Asn Leu Val Ser Glu Ala Gly Leu Thr Gln Lys His  
 1 5 10 15  
 35 Leu Ile His Gly Asp Lys Glu Leu Phe Gln His Glu Leu Lys Thr Ile  
 20 25 30  
 Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile Pro Ser  
 35 40 45  
 Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val Ile Val  
 40 50 55 60



Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe Leu Asn Val Cys Arg  
 65 70 75 80  
 His Arg Gly Lys Thr Leu Val His Ala Glu Ala Gly Asn Ala Lys Gly  
 85 90 95  
 5 Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly Glu Leu  
 100 105 110  
 Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Asp Thr Ile Lys Lys  
 115 120 125  
 Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile Glu Ser Phe His Gly  
 10 130 135 140  
 Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro Thr Leu Val Asp Tyr  
 145 150 155 160  
 Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His Ser Gly  
 165 170 175  
 15 Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys Ala Asn  
 180 185 190  
 Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His Val Gly  
 195 200 205  
 Trp Thr His Ala Ser Ser Leu Arg Ser Gly Gln Ser Ile Phe Thr Pro  
 20 210 215 220  
 Leu Ala Gly Asn Ala Met Leu Pro Pro Glu Gly Ala Gly Leu Gln Met  
 225 230 235 240  
 Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly Tyr Ser  
 245 250 255  
 25 Gly Val His Ser Ala Asp Leu Val Pro Glu Met Met Ala Phe Gly Gly  
 260 265 270  
 Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly Asp Val Arg Ala Arg  
 275 280 285  
 Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn Ser Ile  
 30 290 295 300  
 Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp Glu Asn  
 305 310 315 320  
 Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met Pro Glu  
 325 330 335  
 35 Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val Gly Pro  
 340 345 350  
 Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr Glu Ser  
 355 360 365  
 Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp Leu Ile Ala Asn  
 40 370 375 380

Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys Tyr Pro Gly Val  
 385 390 395 400  
 Val Ala Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe Tyr Arg  
 405 410 415  
 5 Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala Glu Phe Glu Asn  
 420 425 430  
 Thr Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr Thr Asp Arg  
 435 440 445

10 <210> 23  
 <211> 447  
 <212> PRT  
 <213> Artificial Sequence

15 <220>  
 <223> A polypeptide encoded by SEQ ID NO:12

<400> 23  
 Met Ser Tyr Gln Asn Leu Val Ser Glu Ala Gly Leu Thr Gln Lys Leu  
 20 1 5 10 15  
 Leu Ile His Gly Asp Lys Glu Leu Phe Gln His Glu Leu Lys Thr Ile  
 20 25 30  
 Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile Pro Ser  
 35 40 45  
 25 Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val Ile Val  
 50 55 60  
 Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe Leu Asn Val Cys Arg  
 65 70 75 80  
 His Arg Gly Lys Thr Leu Val His Thr Glu Ala Gly Asn Ala Lys Gly  
 30 85 90 95  
 Phe Val Cys Gly Tyr His Gly Trp Gly Tyr Gly Ser Asn Gly Glu Leu  
 100 105 110  
 Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Asp Ala Ile Lys Lys  
 115 120 125  
 35 Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile Glu Ser Phe His Gly  
 130 135 140  
 Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro Pro Leu Ile Asp Tyr  
 145 150 155 160  
 Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Thr Phe Lys His Ser Gly  
 40 165 170 175

Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Val Lys Ala Asn  
 180 185 190  
 Trp Lys Pro Phe Ala Glu Asn Phe Val Gly Asp Ile Tyr His Val Gly  
 195 200 205  
 5 Trp Thr His Ala Ala Ala Leu Arg Ala Gly Gln Ser Val Phe Ser Ser  
 210 215 220  
 Leu Ala Gly Asn Ala Lys Leu Pro Pro Glu Gly Ala Gly Leu Gln Met  
 225 230 235 240  
 Thr Ser Lys Tyr Gly Ser Gly Met Gly Leu Thr Trp Asp Tyr Tyr Ser  
 10 245 250 255  
 Gly Asn Phe Ser Ala Asp Met Val Pro Asp Leu Met Ala Phe Gly Ala  
 260 265 270  
 Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly Asp Val Arg Ala Arg  
 275 280 285  
 15 Ile Tyr Arg Ser Ile Leu Asn Gly Thr Val Phe Pro Asn Asn Ser Phe  
 290 295 300  
 Leu Thr Gly Ser Ala Thr Phe Lys Val Trp Asn Pro Ile Asp Glu Asn  
 305 310 315 320  
 Thr Thr Glu Val Trp Thr Tyr Ala Phe Val Glu Lys Asp Met Pro Glu  
 20 325 330 335  
 Asp Leu Lys Arg Arg Leu Ala Asp Ala Ala Gln Arg Ser Val Gly Pro  
 340 345 350  
 Ala Gly Phe Trp Glu Ser Asp Asp Asn Glu Asn Met Glu Thr Leu Ser  
 355 360 365  
 25 Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp Gln Ile Ala Ser  
 370 375 380  
 Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys Tyr Pro Gly Val  
 385 390 395 400  
 Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe Tyr Arg  
 30 405 410 415  
 Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala Glu Phe Glu Asn  
 420 425 430  
 Ala Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr Thr Asp Arg  
 435 440 445

35

&lt;210&gt; 24

&lt;211&gt; 451

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

40

&lt;220&gt;

&lt;223&gt; A polypeptide encoded by SEQ ID NO:13

&lt;400&gt; 24

```

5 Met Arg Gln Ala Ile Met Ser Tyr Gln Asn Leu Val Ser Glu Ala Gly
  1             5             10             15
Leu Thr Gln Lys His Leu Ile Tyr Gly Asp Lys Glu Leu Phe Gln His
      20             25             30
Glu Leu Lys Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp
10      35             40             45
Ser Leu Ile Pro Ser Pro Gly Asp Tyr Val Lys Ala Lys Met Gly Val
      50             55             60
Asp Glu Val Ile Val Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe
65             70             75             80
15 Leu Asn Val Cys Arg His Arg Gly Lys Thr Ile Val Asp Ala Glu Ala
      85             90             95
Gly Asn Ala Lys Gly Phe Val Cys Gly Tyr His Gly Trp Gly Tyr Gly
      100            105            110
Ser Asn Gly Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly
20      115            120            125
Asp Ala Ile Lys Lys Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile
      130            135            140
Glu Ser Phe His Gly Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro
145            150            155            160
25 Pro Leu Ile Asp Tyr Leu Gly Asp Val Ala Trp Tyr Leu Glu Pro Thr
      165            170            175
Phe Lys His Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Ala Lys Val
      180            185            190
Val Val Lys Gly Asn Trp Lys Val Phe Ala Glu Asn Phe Val Gly Asp
30      195            200            205
Ile Tyr His Ile Gly Trp Thr His Ala Ser Ile Leu Arg Ala Gly Gln
      210            215            220
Ala Ile Phe Ala Pro Leu Ala Gly Asn Ala Met Leu Pro Pro Glu Gly
225            230            235            240
35 Thr Gly Leu Gln Ala Thr Thr Lys Tyr Gly Ser Gly Ile Gly Val Ser
      245            250            255
Leu Asp Ala Tyr Ser Gly Val Gln Ser Ala Asp Leu Val Pro Glu Met
      260            265            270
Met Ala Phe Gly Gly Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly
40      275            280            285

```

Asp Val Arg Ala Arg Ile Tyr Arg Ser Gln Val Asn Gly Thr Val Phe  
 290 295 300  
 Pro Asn Asn Cys Phe Leu Thr Gly Ala Gly Val Phe Lys Val Phe Asn  
 305 310 315 320  
 5 Pro Ile Asp Glu Asn Thr Thr Glu Ala Trp Thr Tyr Ala Ile Val Glu  
 325 330 335  
 Lys Asp Met Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Ala Gln  
 340 345 350  
 Arg Ser Val Gly Pro Ala Gly Tyr Trp Glu Ser Asp Asp Asn Asp Asn  
 10 355 360 365  
 Met Val Leu Ser Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp  
 370 375 380  
 Leu Ile Ala Asp Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys  
 385 390 395 400  
 15 Tyr Pro Gly Val Val Ser Lys Ser Ala Phe Ser Glu Thr Asn His Arg  
 405 410 415  
 Gly Phe Tyr Arg Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala  
 420 425 430  
 Glu Phe Glu Asn Thr Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr  
 20 435 440 445  
 Thr Asp Arg  
 450

<210> 25  
 25 <211> 2265  
 <212> DNA  
 <213> Pseudomonas sp.

<400> 25  
 30 gagggtagag aaatcgaatg ccccttgcat caaggctcggg ttgacgtttg cacaggcaaa 60  
 gccctgtgag caccctgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120  
 cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt 180  
 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240  
 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca 300  
 35 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360  
 tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga 420  
 gaacgacggg tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggg 480  
 gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg 540  
 ctccaacggg gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa 600  
 40 taaaaaatgt ctgggggtga aagaagtcgc tcgctgggag agcttccatg gcttcatcta 660

cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg 720  
 cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780  
 tgtgatcaag gccaaactga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840  
 gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900  
 5 caatgcccgc ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg 960  
 catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020  
 gatggcattc ggaggcgcaa agcaggaaa gctgaacaaa gaaattggcg atgttcgcgc 1080  
 tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140  
 ctccgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200  
 10 cgccattgtc gaaaaagaca tgccctgagga tctcaagcgc cgcttgggccg actctgttca 1260  
 gcgaacgttc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320  
 ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380  
 cggtgaggac gtatacggcg acgcggctta tccaggcgtc gtcggcaaat cggcgatcgg 1440  
 cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cactgcagca gctccaactg 1500  
 15 ggctgagttc gagcatgctt ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560  
 ctaacagacg agtcgacct gatgatcaat attcaagaag acaagctggt ttccgcccac 1620  
 gacgcgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680  
 actacgctgc tgaccagga agcgcatctg ttggacattc aggccttaccg tgcttggtta 1740  
 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cagcggaact gcgcgcagct 1800  
 20 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860  
 ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920  
 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980  
 cttcacatcc gctccaactg cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040  
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100  
 25 cagcgattcg tcgattacc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160  
 tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220  
 aatgtacgtg tatgggcaat caacaagtcg ttctgataac cggtg 2265

<210> 26

30 <211> 449

<212> PRT

<213> *Pseudomonas* sp.

<400> 26

35 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln

1 5 10 15

Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys

20 25 30

Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile

40 35 40 45

Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 50 55 60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65 70 75 80  
 5 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
 85 90 95  
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 10 115 120 125  
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 130 135 140  
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145 150 155 160  
 15 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
 165 170 175  
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
 180 185 190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 20 195 200 205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
 210 215 220  
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
 225 230 235 240  
 25 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
 245 250 255  
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 260 265 270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 30 275 280 285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 290 295 300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305 310 315 320  
 35 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
 325 330 335  
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Phe  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 40 355 360 365

This document contains  
 information that is  
 exempt from release  
 under the provisions of  
 the Freedom of Information Act

Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 5 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 10 435 440 445  
 Arg

<210> 27

15 <211> 2265

<212> DNA

<213> Artificial Sequence

<220>

20 <223> A modified DNA sequence

<400> 27

	gagggtagag aaatcgaatg ccccttgcac caaggctcgt ttgacgtttg cacaggcaaa	60
	gccctgtgag caccgtgac acagaacatc aaaacatc cagtcaagat tgagaacctg	120
25	cgcgtaataa ttgatttgag ctaagaattt taacaggagg caccgccggc cctagagcgt	180
	aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt	240
	aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca	300
	tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc	360
	tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctccccgca	420
30	gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt	480
	gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg	540
	ctccaacggt gaactgcaga gcgttcatt tgaaaaagat ctgtacggcg agtcgctcaa	600
	taaaaaatgt ctgggggttg aagaagtcgc tcgcgtggag agcttccatg gcttcatcta	660
	cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg	720
35	cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaagg	780
	tgtgatcaag gccaaactga aggcacccgc ggaaaacttt gtgggagatg cataccacgt	840
	gggttgagc cagcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg	900
	caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg	960
	catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt	1020
40	gatggcattc ggaggcgcaa agcaggaaa gctgaacaaa gaaattggcg atgttcgcgc	1080



tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140  
 ctccgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200  
 cgcattgttc gaaaaagaca tgccctgagga tctcaagcgc cgcttgggccg actctgttca 1260  
 gcgaacgggc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320  
 5 ttccgaaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380  
 cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440  
 cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cagtcagca gctccaactg 1500  
 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560  
 ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620  
 10 gacgcggaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680  
 actacgctgc tgaccaggga agcgcatttg ttggacattc aggcctaccg tgcttggtta 1740  
 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cagcggaact gcgcgcagct 1800  
 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860  
 ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920  
 15 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980  
 ctccacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgac 2040  
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100  
 cagcgattcg tcgattaccc agagcgcata ctccagacgc acaatctgat ggtctttctg 2160  
 tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220  
 20 aatgtacgtg tatgggcaat caacaagtcg ttccgataac cgggtg 2265

<210> 28

<211> 2265

<212> DNA

25 <213> Artificial Sequence

<220>

<223> A modified DNA sequence

30 <400> 28

gagggtagag aaatcgaatg ccccttgcat caaggctcgt ttgacgtttg cacaggcaaa 60  
 gccctgtgcg caccctgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120  
 cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccocgggc cctagagcgt 180  
 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240  
 35 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300  
 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360  
 tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga 420  
 gaacgacggt tcgattcgtg cttttctgaa cgtttgcggc catcgtggca agacgctggt 480  
 gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg 540  
 40 ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa 600

taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta 660  
 cgggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg 720  
 cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780  
 tgtgatcaag gccaaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840  
 5 ggggttgacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900  
 caatgcggcg ctaccacctg aaggcgacgg cttgcaaatg acctccaaat acggcagcgg 960  
 catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020  
 gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc 1080  
 tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgcagacctg 1140  
 10 ctcggtgtgt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200  
 cgccattgtc gaaaaagaca tgcttgagga tctcaagcgc cgcttgggcg actctgttca 1260  
 gcgaacggcc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320  
 ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380  
 cggtgaggac gtatacggcg acgcggtcta tcaggcgctc gtcggcaaat cggcgatcgg 1440  
 15 cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaaactg 1500  
 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560  
 ctaacagacg agtcgacct gatgatcaat attcaagaag acaagctggt ttccgcccac 1620  
 gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680  
 actacgctgc tgaccagga agcgcatctt ttggacattc aggcctaccg tgcttggtta 1740  
 20 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800  
 tcagagcgct gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860  
 ctgaaagtcc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920  
 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980  
 cttcacatcc gctccaactg cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040  
 25 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100  
 cagcgattcg tcgattacct agagcgcata cttcagacgc acaatctgat ggtctttctg 2160  
 tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220  
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

30 <210> 29  
 <211> 2265  
 <212> DNA  
 <213> Artificial Sequence

35 <220>  
 <223> A modified DNA sequence

<400> 29  
 gagggtagag aaatcgaatg ccccttgcac caaggctcgt ttgacgtttg cacaggcaaa 60  
 40 gccctgtgcg caccctgtac acagaacatc aaaacatata cagtcaagat tgagaacctg 120

	cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccocgggc cctagagcgt	180
	aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt	240
	aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca	300
	tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc	360
5	tgcccccggc gactatgtta cgcacaaaat ggggattgac gaggtcatcg tctcccgga	420
	gaacgacggg tgcattcgtg cttttctgaa cgtttgcccg catcgtggca agacgctggg	480
	gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg	540
	ctccaacggg gaactgcaga gcgttcatt tgaaaaagat ctgtacggcg agtcgctcaa	600
	taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta	660
10	cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg	720
	cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaagg	780
	tgtgatcaag gccaaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt	840
	gggttggaag cagcgcgtct cgcttcgctc gggggagctc atcttctcgt cgctcgtgg	900
	caatgcggcg ctaccacctg aaggcgcagg cttgcaaag acctccaaat acggcagcgg	960
15	catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt	1020
	gatggcattc ggaggcgcaa agcaggaaaag gctgaacaaa gaaattggcg atgttcgcgc	1080
	tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgcctgacctg	1140
	ctcgggtgtt ttcaaagtat ggaacccgat cgacgcacaa accaccgagg tctggacctg	1200
	cgcattgtc gaaaaagaca tgctgagga tctcaagcgc cgcttgcccg actctgttca	1260
20	gcgaacgacc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc	1320
	ttcgcaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt	1380
	cggtaggagc gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg	1440
	cagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg	1500
	ggctgagttc gagcatgcct ctactacttg gcatactgaa cttacgaaga ctactgatcg	1560
25	ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac	1620
	gacgcgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc	1680
	actacgctgc tgaccagga agcgcatttg ttggacattc aggcctaccg tgcttggtta	1740
	gagcactgcg tggggtcaga ggtgcaatat caggctattt cacgcgaact gcgcgcagct	1800
	tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa	1860
30	ctgaaagtgc gagttgagca tcaactggat cgcacaaact ggggcaacag cccgaagctg	1920
	cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta	1980
	cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc	2040
	ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattgggc	2100
	cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg	2160
35	tgattcagtg accattttta caaatgggtc ctgcaaccgc ggtcaccatt aatcaaaggg	2220
	aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg	2265

&lt;210&gt; 30

&lt;211&gt; 2265

40

&lt;212&gt; DNA

## &lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

5

&lt;400&gt; 30

	gagggtagag aaatcgaatg ccccttgcac caaggtcggg ttgacgtttg cacaggcaaa	60
	gccctgtgcg caccctgcac acagaacatc aaaacatatc cagtcaagat tgagaacctg	120
	cgcgtaaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt	180
10	aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt	240
	aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca	300
	tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc	360
	tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga	420
	gaacgacggg tcgattcgtg cttttctgaa cgtttgcggg catcgtggca agacgctggg	480
15	gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg	540
	ctccaacggg gaactgcaga gcgttcatt tgaaaaagat ctgtacggcg agtcgctcaa	600
	taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta	660
	cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg	720
	cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggg	780
20	tgtgatcaag gccaaactga aggcaccgcg ggaaaacttt gtgggagatg cataccacgt	840
	gggttgagcg cagcgtctct cgttcgctc gggggaggtc atcttctcgt cgctcgctgg	900
	caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaat acggcagcgg	960
	catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt	1020
	gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc	1080
25	tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg	1140
	ctcgggtgtt ttcaaagtat ggaaccgat cgacgcaaac accaccgagg tctggacctg	1200
	cgccattgtc gaaaaagaca tgctgagga tctcaagcgc cgcttgccg actctgttca	1260
	gcgaacgctc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc	1320
	ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt	1380
30	cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg	1440
	cgagaccagt tatcgtgggt tctaccgggc ttaccaggca cacgtcagca gctccaactg	1500
	ggctgagttc gagcatgcct ctactacttg gcatactgaa cttacgaaga ctactgatcg	1560
	ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggg ttccgcccac	1620
	gacgcgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc	1680
35	actacgtgc tgaccagga agcgcatttg ttggacattc aggtttaccg tgcttggtta	1740
	gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct	1800
	tcagagcgct gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa	1860
	ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg	1920
	cgttttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta	1980
40	cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc	2040

ttctacgccg cccggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100  
 cagcgattcg tgcattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160  
 tgattcagtg accattttta caaatggta ctgcaaccgc ggtcaccatt aatcaaaggg 2220  
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

5

&lt;210&gt; 31

&lt;211&gt; 2265

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

10

&lt;220&gt;

&lt;223&gt; A modified DNA sequence

&lt;400&gt; 31

15 gagggtagag aaatcgaatg ccccttgcac caaggtcggg ttgacgtttg cacaggcaaa 60  
 gccctgtgcg caccgtgac acagaacatc aaaacatata cagtcaagat tgagaacctg 120  
 cgcgtaatga ttgatttgag ctaagaattt taacaggagg cccccgggc cctagagcgt 180  
 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240  
 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca 300  
 20 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360  
 tgcccccggc gactatgtta cgcgaaaaat ggggattgac gaggtcatcg tctcccgga 420  
 gaacgacggt tgcattcgtg cttttctgaa cgtttgcgg catcgtggca agacgtggt 480  
 gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg 540  
 ctccaacggt gaactgcaga gcgttcatt tgaaaaagat ctgtacggcg agtcgctcaa 600  
 25 taataaatgt ctgggggtta aagaagtcgc tcgcgtggag agcttccatg gcttcatcta 660  
 cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggt 720  
 cctggaacct atgttcaagc attccggcgg tttagaactg gtcggctctc caggcaagg 780  
 tgtgatcaag gccaaactga aggcaccgcg gaaaaacttt gtgggagatg cataccacgt 840  
 ggggttgacg cagcgtctt cgttcgctc gggggagtct atcttctcgt cgctcgctgg 900  
 30 caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg 960  
 catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020  
 gatggcattc ggaggcgcaa agcaggaaa gctgaacaaa gaaattggcg atgttcgcgc 1080  
 tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140  
 ctgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacct 1200  
 35 cgccattgtc gaaaaagaca tgccctagga tctcaagcgc cgcttgggcg actctgttca 1260  
 gcgaacgacg gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320  
 ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380  
 cgggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440  
 cgagaccagt tctcgtggtt tctaccgggc ttaccaggca cagtcagca gctccaactg 1500  
 40 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560

```

ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgaccagga agcgcatctg ttggacattc aggottaccg tgcttggtta 1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
5 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgct 2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggct 2100
10 cagcgattcg tcgattaccg agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

```

<210> 32

15 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

20 <223> A polypeptide encoded by SEQ ID NO:27

<400> 32

```

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1             5             10             15
25 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
    20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
    35             40             45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
30    50             55             60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65             70             75             80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
    85             90             95
35 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
    100            105            110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
    115            120            125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
40    130            135            140

```

His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145 150 155 160  
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
 165 170 175  
 5 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
 180 185 190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 195 200 205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
 10 210 215 220  
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
 225 230 235 240  
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
 245 250 255  
 15 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 260 265 270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 275 280 285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 20 290 295 300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305 310 315 320  
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
 325 330 335  
 25 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Gly  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 355 360 365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 30 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 35 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 435 440 445

Arg

<210> 33  
 <211> 449  
 <212> PRT  
 <213> Artificial Sequence

5

<220>  
 <223> A polypeptide encoded by SEQ ID NO:28

<400> 33

```

10 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
    1             5             10             15
    Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
        20             25             30
    Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
15      35             40             45
    Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
        50             55             60
    Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
    65             70             75             80
20 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
        85             90             95
    Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
        100            105            110
    Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
25      115            120            125
    Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
        130            135            140
    His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
    145            150            155            160
30 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
        165            170            175
    Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
        180            185            190
    Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
35      195            200            205
    Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
        210            215            220
    Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
    225            230            235            240
40 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
  
```



245 250 255  
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 260 265 270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 5 275 280 285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 290 295 300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305 310 315 320  
 10 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
 325 330 335  
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Ala  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 15 355 360 365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 20 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 25 435 440 445  
 Arg

<210> 34

30 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

35 <223> A polypeptide encoded by SEQ ID NO:29

<400> 34

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln  
 1 5 10 15  
 40 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys

	20		25		30
	Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile				
	35		40		45
	Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val				
5	50		55		60
	Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val				
	65		70		75
	Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala				
		85		90	95
10	Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly				
	100		105		110
	Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu				
	115		120		125
	Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe				
15	130		135		140
	His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met				
	145		150		155
	Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His				
		165		170	175
20	Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys				
	180		185		190
	Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His				
	195		200		205
	Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe				
25	210		215		220
	Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu				
	225		230		235
	Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly				
		245		250	255
30	Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe				
	260		265		270
	Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg				
	275		280		285
	Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn				
35	290		295		300
	Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp				
	305		310		315
	Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met				
		325		330	335
40	Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Thr				

340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 355 360 365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 5 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 10 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 435 440 445  
 Arg  
 15  
 <210> 35  
 <211> 449  
 <212> PRT  
 20 <213> Artificial Sequence  
 <220>  
 <223> A polypeptide encoded by SEQ ID NO:30  
 25 <400> 35  
 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln  
 1 5 10 15  
 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys  
 20 25 30  
 30 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 35 40 45  
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 50 55 60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 35 65 70 75 80  
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
 85 90 95  
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 40 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu

	115	120	125
	Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe		
	130	135	140
	His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met		
5	145	150	155
	Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His		160
	165	170	175
	Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys		
	180	185	190
10	Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His		
	195	200	205
	Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe		
	210	215	220
	Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu		
15	225	230	235
	Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly		240
	245	250	255
	Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe		
	260	265	270
20	Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg		
	275	280	285
	Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn		
	290	295	300
	Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp		
25	305	310	315
	Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met		320
	325	330	335
	Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Leu		
	340	345	350
30	Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr		
	355	360	365
	Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu		
	370	375	380
	Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro		
35	385	390	395
	Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe		400
	405	410	415
	Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe		
	420	425	430
40	Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp		

435 440 445

Arg

5 <210> 36  
 <211> 449  
 <212> PRT  
 <213> Artificial Sequence

10 <220>  
 <223> A polypeptide encoded by SEQ ID NO:31

<400> 36

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln  
 15 1 5 10 15  
 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys  
 20 25 30  
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 35 40 45  
 20 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 50 55 60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65 70 75 80  
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
 25 85 90 95  
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 115 120 125  
 30 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 130 135 140  
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145 150 155 160  
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
 35 165 170 175  
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
 180 185 190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 195 200 205  
 40 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe

	210	215	220
	Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu		
	225	230	235 240
	Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly		
5	245	250	255
	Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe		
	260	265	270
	Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg		
	275	280	285
10	Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn		
	290	295	300
	Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp		
	305	310	315 320
	Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met		
15	325	330	335
	Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Ile		
	340	345	350
	Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr		
	355	360	365
20	Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu		
	370	375	380
	Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro		
	385	390	395 400
	Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe		
25	405	410	415
	Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe		
	420	425	430
	Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp		
	435	440	445

30 Arg

&lt;210&gt; 37

&lt;211&gt; 22

35 &lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; An oligonucleotide

40

<400> 37  
 ttcagcgaac ggtcgggcct gc 22

5      <210> 38  
         <211> 27  
         <212> DNA  
         <213> Artificial Sequence

10      <220>  
         <223> An oligonucleotide

<400> 38  
 gccgggcctc ttgcgggata tcgtcca 27

15      <210> 39  
         <211> 27  
         <212> DNA  
         <213> Artificial Sequence

20      <220>  
         <223> An oligonucleotide

<400> 39  
 gttgccattg ctgcaggcat cgtggtg 27

25      <210> 40  
         <211> 33  
         <212> DNA  
         <213> Artificial Sequence

30      <220>  
         <223> An oligonucleotide

<400> 40  
 35 gaggcacccg cggaagcttt tgtgggagat gca 33

<210> 41  
 <211> 30  
 <212> DNA  
 40      <213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; An oligonucleotide

&lt;400&gt; 41

5 gcacccgcgg aacaatttgt gggagatgca

30

&lt;210&gt; 42

&lt;211&gt; 21

&lt;212&gt; DNA

10 &lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; An oligonucleotide

15 &lt;400&gt; 42

ccgcggaaag ctttgtggga g

21

&lt;210&gt; 43

&lt;211&gt; 24

20 &lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; An oligonucleotide

25

&lt;400&gt; 43

ccgcggaaaa gcttgtggga gatg

24

&lt;210&gt; 44

30 &lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

35 &lt;223&gt; An oligonucleotide

&lt;400&gt; 44

cgcggaac gttgtgggag atg

23

40 &lt;210&gt; 45



<213> Artificial Sequence

<223> An oligonucleotide

atattcaggt gcgcatagcg cag

23

<213> Artificial Sequence

<223> An oligonucleotide

20 ggacggatat tcagggctcc atagcgcaga cttg

34

25           <213> Artificial Sequence

<223> An oligonucleotide

gacggatatt caggtaacca tagcgcagac ttg

33

<213> Artificial Sequence

<223> An oligonucleotide

40

	<400> 48	
	ggtgttttca aagtcgcgaa cccgatcgac	30
	<210> 49	
5	<211> 26	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
10	<223> An oligonucleotide	
	<400> 49	
	ctgttcagcg aaacttcggg cctgct	26
15	<210> 50	
	<211> 26	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> An oligonucleotide	
	<400> 50	
	ctgttcagcg aaggttcggg cctgct	26
25	<210> 51	
	<211> 26	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> An oligonucleotide	
	<400> 51	
35	ctgttcagcg aagcttcggg cctgct	26
	<210> 52	
	<211> 22	
	<212> DNA	
40	<213> Artificial Sequence	

<223> An oligonucleotide

22

10           <213> Artificial Sequence

<211> 30

<213> Artificial Sequence

<223> An oligonucleotide

30

<210> 54

20                   <212> DNA

<220>

25

21

<210> 55

<212> DNA

<220>

<400> 55

30

40                    <210> 56

&lt;211&gt; 2265

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

5 &lt;220&gt;

&lt;223&gt; A modified DNA sequence

&lt;400&gt; 56

	gagggtagag aaatcgaatg ccccttgcac caaggtcggg ttgacgtttg cacaggcaaa	60
10	gcctgtgag caccctgac acagaacac aaaacatata cagtcaagat tgagaacctg	120
	cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt	180
	aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt	240
	aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca	300
	tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc	360
15	tgccccggc gactatgta ccgcaaaaat ggggattgac gaggtcatcg tctccggca	420
	gaacgacggg tcgattcgtg cttttctgaa cgtttgccg catcgtggca agacgctggg	480
	gagcgtggaa gccggcaatg ccaaggttt tgttgcagc tatcacggc ggggcttcgg	540
	ctccaacggg gaactgcaga gcgttcatt tgaaaagat ctgtacggcg agtcgctcaa	600
	taaaaaatgt ctgggggtga aagaagtcgc tcgctggag agcttccatg gcttcatcta	660
20	cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg	720
	cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaagg	780
	tgtgatcaag gccaaactga aggcacccgc ggaaaacttt gtgggagatg cataccacgt	840
	gggttgagc cagcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg	900
	caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg	960
25	catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt	1020
	gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgtcgcgc	1080
	tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg	1140
	ctcgggtgtt ttcaaagtat ggaaccgat cgacgcaaac accaccgagg tctggacctg	1200
	cgccattgtc gaaaaagaca tgcttgagga tctcaagcgc cgcttgcccg actctgttca	1260
30	gcgaacgtgg gggcctgctg gcttctggga aagcgacgac aatgacaata tggaacagc	1320
	ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt	1380
	cggtagaggc gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg	1440
	cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cagtcagca gctccaactg	1500
	ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg	1560
35	ctaacagacg agtcgacct gatgatcaat attcaagaag acaagctggg ttccgcccac	1620
	gacgccgaag agattcttgc tttcttcaat tgccacgact ctgctttgca acaagaagcc	1680
	actacgtgc tgaccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta	1740
	gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgagct	1800
	tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa	1860
40	ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg	1920

cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaataa caaagagcta 1980  
 cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040  
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100  
 cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160  
 5 tgattcagt accattttta caaatggta ctgcaaccgc ggtcaccatt aatcaaagg 2220  
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

<210> 57

<211> 2265

10 <212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

15

<400> 57

gagggtagag aaatcgaatg ccccttgcac caaggctcgt ttgacgtttg cacaggcaaa 60  
 gccctgtgcg caccctgtac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120  
 cgcgtaatga ttgatttgag ctaagaatct taacaggagg caccctgggc cctagagcgt 180  
 20 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggg 240  
 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac tttccaaca 300  
 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360  
 tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga 420  
 gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggt 480  
 25 gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcaccggt ggggcttcgg 540  
 ctccaacggt gaactgcaga gcgttccatt tgaaaagat ctgtacggcg agtcgctcaa 600  
 taaaaaatgt ctgggggtga aagaagtgcg tcgcgtggag agcttccatg gcttcacta 660  
 cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg 720  
 cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaagg 780  
 30 tgtgatcaag gcccaactga aggcaccgcg ggaaaacttt gtgggagatg cataccacgt 840  
 ggggttgacg cagcgtctt cgcttcgctc gggggagtct atcttctcgt cgtcgtctgg 900  
 caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg 960  
 catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020  
 gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc 1080  
 35 tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgetgacctg 1140  
 ctgggtgtt ttcaaagtat ggaaccgat cgacgcaaac accaccgagg tctggacctg 1200  
 cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttgcccg actctgttca 1260  
 gcgaacggtc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320  
 ttcgcaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380  
 40 cggtaggac gtatacggcg acgcgtgcta tccaggcgtc gtcggcaaat cggcgatcgg 1440

cgagaccagt tatcgtgggt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500  
 ggctgagttc gagcatgect ctactacttg gcatactgaa cttacgaaga ctactgatcg 1560  
 ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggg ttcgcccac 1620  
 gagccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680  
 5 actacgctgc tgaccaggga agcgcatctt ttggacattc aggcttaccg tgcttggtta 1740  
 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cagcggaact gcgcgcagct 1800  
 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860  
 ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920  
 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980  
 10 cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgct 2040  
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggct 2100  
 cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160  
 tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220  
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

15

&lt;210&gt; 58

&lt;211&gt; 449

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

20

&lt;220&gt;

&lt;223&gt; A polypeptide encoded by SEQ ID NO:56

&lt;400&gt; 58

25 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln  
 1 5 10 15  
 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys  
 20 25 30  
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 30 35 40 45  
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 50 55 60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65 70 75 80  
 35 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
 85 90 95  
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 40 115 120 125

Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 130 135 140  
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145 150 155 160  
 5 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
 165 170 175  
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
 180 185 190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 10 195 200 205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
 210 215 220  
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
 225 230 235 240  
 15 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
 245 250 255  
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 260 265 270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 20 275 280 285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 290 295 300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305 310 315 320  
 25 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
 325 330 335  
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Trp  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 30 355 360 365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 35 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 40 435 440 445

Arg

<210> 59  
 5 <211> 449  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 10 <223> A polypeptide encoded by SEQ ID NO:57  
  
 <400> 59  
 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln  
 1 5 10 15  
 15 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys  
 20 25 30  
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile  
 35 40 45  
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val  
 20 50 55 60  
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val  
 65 70 75 80  
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala  
 85 90 95  
 25 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly  
 100 105 110  
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu  
 115 120 125  
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe  
 30 130 135 140  
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met  
 145 150 155 160  
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His  
 165 170 175  
 35 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys  
 180 185 190  
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His  
 195 200 205  
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe  
 40 210 215 220



Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu  
 225 230 235 240  
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly  
 245 250 255  
 5 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe  
 260 265 270  
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg  
 275 280 285  
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn  
 10 290 295 300  
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp  
 305 310 315 320  
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met  
 325 330 335  
 15 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Tyr  
 340 345 350  
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr  
 355 360 365  
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu  
 20 370 375 380  
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro  
 385 390 395 400  
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe  
 405 410 415  
 25 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe  
 420 425 430  
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp  
 435 440 445

Arg

30

&lt;210&gt; 60

&lt;211&gt; 26

&lt;212&gt; DNA

35

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; An oligonucleotide

40

&lt;400&gt; 60

gttcagcgaa cgggcgggcc tgctgg

26

<210> 61

<211> 26

5 <212> DNA

<213> Artificial Sequence

<220>

<223> An oligonucleotide

10

<400> 61

gttcagcgaa cggccgggcc tgctgg

26

<210> 62

15 <211> 26

<212> DNA

<213> Artificial Sequence

<220>

20 <223> An oligonucleotide

<400> 62

gttcagcgaa cgaccgggcc tgctgg

26

25 <210> 63

<211> 26

<212> DNA

<213> Artificial Sequence

30 <220>

<223> An oligonucleotide

<400> 63

gttcagcgaa cgatcgggcc tgctgg

26

35

<210> 64

<211> 26

<212> DNA

<213> Artificial Sequence

40

<223> An oligonucleotide

26

5 gttcagcgaa cgtggggggcc tgctgg

<211> 25

10

<213> Artificial Sequence

<223> An oligonucleotide

15

<400> 65

ttcagcgaac gtacgggcct gctgg

25